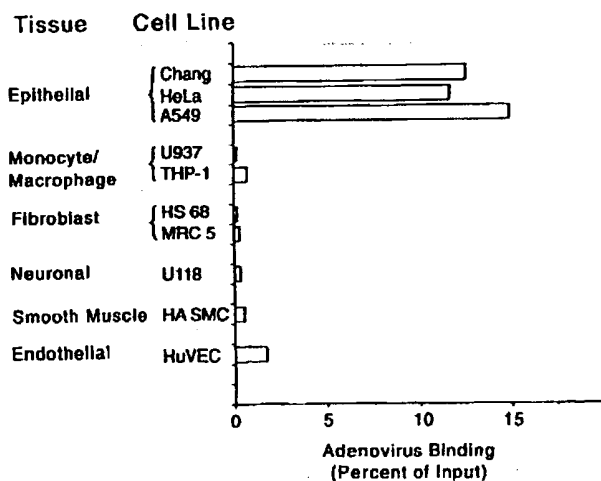




INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/34, C07K 14/075, C12N 15/86, 5/10		A2	(11) International Publication Number: WO 97/20051
			(43) International Publication Date: 5 June 1997 (05.06.97)
(21) International Application Number: PCT/US96/19150		KOVESDI, Imre [CA/US]; 206 Rollins Avenue, Rockville, MD 20852 (US). BROUGH, Douglas, E. [US/US]; 3900 Shallowbrook Lane, Olney, MD 20832 (US).	
(22) International Filing Date: 27 November 1996 (27.11.96)		(74) Agents: KILYK, John, Jr. et al.; Leydig, Voit & Mayer, Ltd., Two Prudential Plaza, Suite 4900, 180 North Stetson, Chicago, IL 60601-6780 (US).	
(30) Priority Data: 08/563,368 28 November 1995 (28.11.95) US 08/700,846 21 August 1996 (21.08.96) US 08/701,124 21 August 1996 (21.08.96) US		(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(60) Parent Applications or Grants (63) Related by Continuation US 08/700,846 (CIP) Filed on 21 August 1996 (21.08.96) US 08/701,124 (CIP) Filed on 21 August 1996 (21.08.96) US 08/563,368 (CIP) Filed on 28 November 1995 (28.11.95)		Published <i>Without international search report and to be republished upon receipt of that report.</i>	
(71) Applicant (for all designated States except US): GENVEC, INC. [US/US]; 12111 Parklawn Drive, Rockville, MD 20852 (US).			
(72) Inventors; and (75) Inventors/Applicants (for US only): WICKHAM, Thomas, J. [US/US]; 5 Charen Court, Potomac, MD 20854 (US).			

(54) Title: VECTORS AND METHODS FOR GENE TRANSFER TO CELLS



(57) Abstract

The present invention provides a chimeric adenovirus coat protein, which differs from the wild-type coat protein by the introduction of a nonnative amino acid sequence. Such a chimeric adenovirus coat protein according to the invention is able to direct entry into cells of a vector comprising the coat protein that is more efficient than entry into cells of a vector that is identical except for comprising a wild-type adenovirus coat protein rather than the chimeric adenovirus coat protein. The chimeric coat protein preferably is a fiber, hexon, or penton protein. The present invention also provides an adenoviral vector that comprises the chimeric adenovirus coat protein, as well as methods of constructing and using such a vector.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

VECTORS AND METHODS FOR GENE TRANSFER TO CELLS

TECHNICAL FIELD OF THE INVENTION

The present invention pertains to a chimeric
5 adenovirus coat protein which is able to direct entry
into cells of a vector comprising the coat protein that
is more efficient than a similar vector having a wild-
type adenovirus coat protein. Such a chimeric coat
protein is a fiber, hexon, or penton protein. The
10 present invention also pertains to a recombinant vector
comprising such a chimeric adenoviral coat protein, and
to methods of constructing and using such a vector.

BACKGROUND OF THE INVENTION

15 Adenoviruses belong to the family Adenoviridae,
which is divided into two genera, namely Mastadenovirus
and Aviadenovirus. Adenoviruses are nonenveloped,
regular icosahedrons of about 65 to 80 nanometers in
diameter (Horne et al., J. Mol. Biol., 1, 84-86 (1959)).
20 The adenoviral capsid is composed of 252 capsomeres of
which 240 are hexons and 12 are pentons (Ginsberg et al.,
Virology, 28, 782-783 (1966)). The hexons and pentons
are derived from three different viral polypeptides
(Maizel et al., Virology, 36, 115-125 (1968); Weber et
25 al., Virology, 76, 709-724 (1977)). The hexon comprises
three identical polypeptides of 967 amino acids each,
namely polypeptide II (Roberts et al., Science, 232,
1148-1151 (1986)). The penton contains a penton base,
which is bound to the capsid, and a fiber, which is
30 noncovalently bound to and projects from the penton base.
The fiber protein comprises three identical polypeptides
of 582 amino acids each, namely polypeptide IV. The
adenovirus serotype 2 (Ad2) penton base protein is a
ring-shaped complex composed of five identical protein
35 subunits of 571 amino acids each, namely polypeptide III
(Boudin et al., Virology, 92, 125-138 (1979)). Proteins
IX, VI, and IIIa are also present in the adenoviral coat

and are thought to stabilize the viral capsid (Stewart et al., Cell, 67, 145-154 (1991); Stewart et al., EMBO J., 12(7), 2589-2599 (1993)).

Once an adenovirus attaches to a cell, it undergoes
5 receptor-mediated internalization into clathrin-coated
endocytic vesicles of the cell (Svensson et al., J. Virol., 51, 687-694 (1984); Chardonnet et al., Virology, 40, 462-477 (1970)). Virions entering the cell undergo a
stepwise disassembly in which many of the viral
10 structural proteins are shed (Greber et al., Cell, 75,
477-486 (1993)). During the uncoating process, the viral
particles cause disruption of the cell endosome by a pH-
dependent mechanism (Fitzgerald et al., Cell, 32, 607-617
(1983)), which is still poorly understood. The viral
15 particles are then transported to the nuclear pore
complex of the cell (Dales et al., Virology, 56, 465-483
(1973)), where the viral genome enters the nucleus, thus
initiating infection.

An adenovirus uses two separate cellular receptors,
20 both of which must be present, to efficiently attach to
and infect a cell (Wickham et al., Cell, 73, 309-319
(1993)). First, the Ad2 fiber protein attaches the virus
to a cell by binding to an as yet unidentified receptor.
Then, the penton base binds to α_v integrins, which are a
25 family of a heterodimeric cell-surface receptors that
mediate cellular adhesion to the extracellular matrix
molecules, as well as other molecules (Hynes, Cell, 69,
11-25 (1992)).

The fiber protein is a trimer (Devaux et al., J. Molec. Biol., 215, 567-588 (1990)) consisting of a tail,
30 a shaft, and a knob. The fiber shaft region is composed
of repeating 15 amino acid motifs, which are believed to
form two alternating β -strands and β -bends (Green et al.,
EMBO J., 2, 1357-1365 (1983)). The overall length of the
35 fiber shaft region and the number of 15 amino-acid
repeats differ between adenoviral serotypes. For
example, the Ad2 fiber shaft is 37 nanometers long and

contains 22 repeats, whereas the Ad3 fiber is 11 nanometers long and contains 6 repeats. The receptor binding domain of the fiber protein is localized in the knob region encoded by the last 200 amino acids of the protein (Henry et al., J. Virology, 68(8), 5239-5246 (1994)). The regions necessary for trimerization are also located in the knob region of the protein (Henry et al. (1994), supra). A deletion mutant lacking the last 40 amino acids does not trimerize and also does not bind to penton base (Novelli et al., Virology, 185, 365-376 (1991)). Thus, trimerization of the fiber protein is necessary for penton base binding. Nuclear localization signals that direct the protein to the nucleus to form viral particles following its synthesis in the cytoplasm are located in the N-terminal region of the protein (Novelli et al. (1991), supra). The fiber, together with the hexon, are the main antigenic determinants of the virus and also determine the serotype specificity of the virus (Watson et al., J. Gen. Virol., 69, 525-535 (1988)).

Recombinant adenoviral vectors have been used for the cell-targeted transfer of one or more recombinant genes to diseased cells or tissue in need of treatment. Such vectors are characterized by the advantage of not requiring host cell proliferation for expression of adenoviral proteins (Horwitz et al., In Virology, Raven Press, New York, vol. 2, pp. 1679-1721 (1990); and Berkner, BioTechniques, 6, 616 (1988)). Moreover, if the targeted tissue for somatic gene therapy is the lung, these vectors have the added advantage of being normally trophic for the respiratory epithelium (Straus, In Adenoviruses, Plenan Press, New York, pp. 451-496 (1984)).

Other advantages of adenoviruses as potential vectors for human gene therapy are: (i) recombination is rarely observed with use of such vectors; (ii) there are no known associations of human malignancies with

adenoviral infections despite common human infection with adenoviruses; (iii) the adenoviral genome (which is a linear, double-stranded DNA) can be manipulated to accommodate foreign genes that range in size; (iv) an
5 adenoviral vector does not insert its DNA into the chromosome of a cell, so its effect is impermanent and unlikely to interfere with the cell's normal function; (v) the adenovirus can infect non-dividing or terminally differentiated cells, such as cells in the brain and
10 lungs; and (vi) live adenovirus, having as an essential characteristic the ability to replicate, has been safely used as a human vaccine (Horwitz et al. (1990), supra; Berkner et al. (1988), supra; Straus et al. (1984), supra; Chanock et al., JAMA, 195, 151 (1966); Haj-Ahmad
15 et al., J. Virol., 57, 267 (1986); and Ballay et al., EMBO, 4, 3861 (1985); PCT patent application WO 94/17832).

A drawback to adenovirus-mediated gene therapy is that significant decreases in gene expression are
20 observed after two weeks following administration of the vector. In many therapeutic applications, the loss of expression requires re-administration of the viral vector. However, following re-administration, neutralizing antibodies are raised against both the fiber
25 and hexon proteins of the viral vector (Wohlfart, J. Virology, 62, 2321-2328 (1988); Wohlfart et al., J. Virology, 56, 896-903 (1985)). This antibody response against the virus can prevent effective re-administration of the viral vector.

30 Another drawback of using recombinant adenovirus in gene therapy is that certain cells are not readily amenable to adenovirus-mediated gene delivery. For instance, lymphocytes, which lack the α_v integrin adenoviral receptors, are impaired in the uptake of
35 adenoviruses (Silver et al., Virology 165, 377-387 (1988); Horvath et al., J. Virology, 62(1), 341-345 (1988)). This lack of ability to infect all cells has

lead researchers to seek out ways to introduce adenovirus into cells that cannot be infected by adenovirus, e.g. due to lack of adenoviral receptors. In particular, the virus can be coupled to a DNA-polylysine complex
5 containing a ligand (e.g., transferrin) for mammalian cells (e.g., Wagner et al., Proc. Natl. Acad. Sci., 89, 6099-6103 (1992); PCT patent application WO 95/26412). Similarly, adenoviral fiber protein can be sterically blocked with antibodies, and tissue-specific antibodies
10 can be chemically linked to the viral particle (Cotten et al., Proc. Natl. Acad. Sci. USA, 89, 6094-6098 (1992)).

However, these approaches are disadvantageous in that they require additional steps that covalently link large molecules, such as polylysine, receptor ligands,
15 and antibodies, to the virus (Cotten (1992), supra; Wagner et al., Proc. Natl. Acad. Sci., 89, 6099-6103 (1992)). This adds to the size of the resultant vector as well as its cost of production. Moreover, the targeted particle complexes are not homogeneous in
20 structure, and their efficiency is sensitive to the relative ratios of viral particles, linking molecules, and targeting molecules used. Thus, this approach for expanding the repertoire of cells amenable to adenoviral-mediated gene therapy is less than optimal.

25 Recently, the efficiency of adenovirus-mediated gene transfer *in vivo* to even those cells which adenovirus has been reputed to enter with high efficiency has been called into question (Grubb et al., Nature, 371, 802-806 (1994); Dupuit et al., Human Gene Therapy, 6,
30 1185-1193 (1995)). The fiber receptor by means of which adenovirus initially contacts cells has not been identified; and its representation in different tissues has not been examined. It is generally assumed that epithelial cells in the lung and gut produce sufficient
35 levels of the fiber receptor to allow their optimal transduction. However, no studies have confirmed this point to date. In fact, studies have suggested that

adenovirus gene delivery to differentiated lung epithelium is less efficient than delivery to proliferating or to undifferentiated cells (Grubb et al., supra; Dupuit et al., supra).

5 Similarly, adenovirus has been shown to transduce a large number of tissues including lung epithelial cells (Rosenfeld et al., Cell, 68, 143-155 (1992)), muscle cells (Quantin et al., Proc. Natl. Acad. Sci., 89, 2581-2584 (1992)), endothelial cells (Lemarchand et al., Proc.
10 Natl. Acad. Sci., 89, 6482-6486 (1992)), fibroblasts (Anton et al., J. Virol., 69, 4600-4606 (1995), and neuronal cells (LaSalle et al., Science, 259, 988-990 (1993)). However, in many of these studies, very high levels of virus particles have been used to achieve
15 transduction, often exceeding 100 plaque forming units (pfu)/cell, and corresponding to a multiplicity of infection (MOI) of 100. The requirement for a high MOI to achieve transduction is disadvantageous inasmuch as any immune response associated with adenoviral infection
20 necessarily would be exacerbated with use of high doses.

Accordingly, there remains a need for vectors, such as adenoviral vectors, that are capable of infecting cells with a high efficiency, especially at lower MOIs, and that demonstrate an increased host cell range of
25 infectivity. The present invention seeks to overcome at least some of the aforesaid problems of recombinant adenoviral gene therapy. In particular, it is an object of the present invention to provide a vector (such as an adenoviral vector) having a broad host range, and an
30 ability to enter cells with a high efficiency, even at a reduced MOI, thereby reducing the amount of recombinant adenoviral vector administered and any side-effects/complications resulting from such administration. A further object of the present invention is to provide a
35 method of gene therapy involving the use of a homogeneous adenovirus, wherein the viral particle is modified at the level of the adenoviral genome, without the need for

additional chemical modifications of viral particles. These and other objects and advantages of the present invention, as well as additional inventive features, will be apparent from the following detailed description.

5

BRIEF SUMMARY OF THE INVENTION

The present invention provides a chimeric adenoviral coat protein (e.g., a fiber, hexon or penton protein), which differs from the wild-type (i.e., native) fiber protein by the introduction of a nonnative amino acid sequence, preferably at or near the carboxyl terminus. The resultant chimeric adenovirus coat protein is able to direct entry into cells of a vector comprising the coat protein that is more efficient than entry into cells of a vector that is identical except for comprising a wild-type adenovirus coat protein rather than the chimeric adenovirus coat protein. One direct result of this increased efficiency of entry is that the chimeric adenovirus coat protein enables the adenovirus to bind to and enter numerous cell types which adenovirus comprising wild-type coat protein typically cannot enter or can enter with only a low efficiency. The present invention also provides an adenoviral vector that comprises the chimeric adenovirus coat protein, and methods of constructing and using such a vector.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a bar graph depicting the binding (percent of input) of wild-type adenovirus to cells derived from different tissues.

Figures 2A-B depict attachment of a nucleic acid sequence at the end of the wild-type adenoviral fiber gene (**Fig. 2A**) to derive a chimeric adenoviral fiber protein (**Fig. 2B**) comprising a nonnative amino acid sequence at the carboxy terminus. As indicated, the length of the polyA tail, and, consequently, the number of lysines in the resultant protein, can vary.

Figure 3 is a schematic diagram depicting the construction of the adenovirus transfer vector containing chimeric fiber protein pAd BS 59-100 UTV by way of intermediary transfer vectors. In particular, pAd NS 83-100 (also known as p193NS 83-100 or pNS 83-100) is used to derive fiber minus (i.e., F⁻) pAd NS 83-100 (also known as p193NS (Δ F) or pNS (Δ F)) (**path A**), pAd NS 83 100 (F⁻) is used to derive pAd NS 83-100 UTV (also known as p193NS (F5⁺), p193 (F5⁺), or pNS (F5⁺)) (**path B**), and pAd NS 83-100 UTV is used to derive pAd BS 59-100 UTV (**path C**).

Figures 4A-D depict the oligonucleotides employed for construction of GV10 UTV, i.e., the primers SEQ ID NO:9 (**Fig. 4A**), SEQ ID NO:10 (**Fig. 4B**), SEQ ID NO:11 (**Fig. 4C**), and SEQ ID NO:12 (**Fig. 4D**).

Figure 5 depicts a Western blot showing the size increase of the chimeric adenoviral fiber protein (UTV) as compared with the wild-type fiber protein (WT).

Figures 6A-B are graphs depicting a comparison of the binding of an adenoviral vector comprising wild-type fiber protein (i.e., GV10, open triangle) and adenoviral vector comprising chimeric fiber protein (i.e., GV10 UTV, filled circle) to a receptor-plus (A549, **Fig. 6A**) and a receptor-minus (HS 68, **Fig. 6B**) cell.

Figure 7 is a graph of UTV bound (counts per minute (CPM)) versus amount of competitor (μ g/ml) for inhibition of binding of chimeric adenoviral fiber protein to receptor-minus cells (i.e., HS 68 fibroblasts) by the soluble factors chondroitin sulfate (open circle); heparin (filled circle); mucin (filled triangle); and salmon sperm DNA (open triangle).

Figure 8 is a graph of UTV bound (CPM) versus enzyme dilution for inhibition of binding of chimeric adenoviral fiber protein to receptor-minus cells (i.e., HS 68 fibroblasts) by the enzymes chondroitinase (open circles, stippled lines); heparinase (open circles, solid lines); and sialidase (triangles, solid lines).

Figure 9 is a bar graph depicting a comparison of transfer of a *lacZ* reporter gene by an adenoviral vector comprising wild-type fiber protein (i.e., GV10) and an adenoviral vector comprising chimeric fiber protein (i.e., GV10 UTV) as assessed by resultant reporter gene expression (i.e., relative light units (RLU)) in various receptor-plus and receptor-minus cells.

Figure 10 is a bar graph depicting a comparison of transfer of a *lacZ* reporter gene by an adenoviral vector comprising wild-type fiber protein (i.e., GV10) and an adenoviral vector comprising chimeric fiber protein (i.e., GV10 UTV) as assessed by resultant reporter expression (i.e., relative light units (RLU)) in mouse lung.

Figure 11 is a bar graph depicting the transfer of a reporter gene (i.e., contained in pGUS) by an adenoviral vector comprising wild-type fiber protein (i.e., GV10, solid bars) and an adenoviral vector comprising chimeric fiber (i.e., GV10 UTV, open bars) potentially bound via a protein/DNA interaction into 293 cells, A549 cells, and H700 T cells.

Figure 12 is a diagram that further depicts the plasmid p193(F5*) (described as pAd NS 83-100 UTV in Figure 3, and also known as p193 (F5*) or pNS (F5*)) used to construct adenovirus fiber chimeras, and the sequence of the C-terminus of the mutated fiber protein present in the plasmid (polyadenylation site emboldened).

Figure 13 is a diagram that depicts plasmid p193NS (F5*) pGS(K7) (also known as p193 (F5*) pGS(K7) or pNS (F5*) pK7) used to construct adenovirus fiber chimeras.

Figure 14 is a diagram that depicts plasmid pBSS 75-100 pGS(null) (also known as pBSS 75-100 Δ E3 pGS(null)).

Figure 15 is a diagram that depicts plasmid pBSS 75-100 pGS(RK32) (also known as pBSS 75-100 Δ E3 pGS(RKKK), or pBSS 75-100 Δ E3 pGS(RKKK2)).

Figure 16 is a diagram that depicts plasmid pBSS 75-100 pGS(RK33) (also known as pBSS 75-100 Δ E3 pGS(RKKK), or pBSS 75-100 Δ E3 pGS(RKKK3)).

Figure 17 is a diagram that depicts plasmid p193NS F5F2K (also known as p193 F5F2K).

Figure 18 is a diagram that depicts plasmid p193NS F5F2K(RKKK), (also known as p193NS F5F2K(RKKK2), p193NS F5F2K(RK32), or p193 F5F2K(RKKK2)).

Figure 19 is a diagram that depicts plasmid p193NS F5F2K(RKKK), (also known as p193NS F5F2K(RKKK3), p193 F5F2K(RKKK3), or p193 F5FK(RK33)).

Figure 20 is a diagram that depicts plasmid pACT (RKKK), (also known as pACT (RKKK3) or pACT (RK33)).

Figure 21 is a diagram that depicts plasmid pACT (RKKK), (also known as pACT (RKKK2) or pACT (RK32)).

Figure 22 is a diagram that depicts plasmid pACT H11.

Figure 23 is a diagram that depicts plasmid pACT H11(RKKK), (also known as pACT H11(RKKK2) or pACT H11(RK32)).

Figure 24 is a diagram that depicts plasmid p193 F5F9sK (also known as p193 F5F9K-Short).

Figure 25 is a diagram that depicts plasmid pSPdelta.

Figure 26 is a diagram that depicts plasmid pSP2alpha. The "j" indicates destroyed Ppu10I sites in the plasmid.

Figure 27 is a diagram that depicts plasmid pSP2alpha2. The "j" indicates destroyed Ppu10I sites in the plasmid.

Figure 28 is a graph of days post-infection versus FFU/cell for 293 cells infected with Ad5 (open circles), AdZ.F(RGD) (closed squares), or AdZ.F(pK7) (open triangles).

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides, among other things, a recombinant adenovirus comprising a chimeric coat protein, such as a chimeric fiber, penton, and/or hexon protein. The chimeric coat protein comprises a nonnative amino acid sequence, in addition to, or in place of, a native amino acid sequence. This nonnative amino acid sequence allows the chimeric fiber (or a vector comprising the chimeric fiber) to more efficiently bind to and enter cells.

Thus, the present invention provides, a chimeric adenovirus coat protein comprising a nonnative amino acid sequence, wherein the coat protein is able to direct entry into cells of a vector comprising the coat protein that is more efficient than entry into cells of a vector that is identical except for comprising a wild-type adenovirus coat protein rather than the chimeric adenovirus coat protein (i.e., in the absence of the chimeric adenovirus coat protein and in the presence of the wild-type adenovirus coat protein).

Chimeric Coat Protein

A "coat protein" according to the invention preferably comprises a fiber protein (especially an adenoviral fiber protein), a penton protein (especially an adenoviral penton protein), and a hexon protein (especially an adenoviral hexon protein). In particular, a coat protein preferably comprises an adenoviral fiber, penton, or hexon protein. Any one of the serotypes of human or nonhuman adenovirus can be used as the source of the coat protein gene, optimally, however, the adenovirus is an Ad2 or Ad5 adenovirus.

The coat protein is "chimeric" in that it comprises amino acid residues that are not typically found in the protein as isolated from wild-type adenovirus (i.e., comprising the native protein, or wild-type protein). The coat protein thus comprises a "nonnative amino acid

sequence". By "nonnative amino acid sequence" is meant any amino acid sequence that is not found in the native fiber of a given serotype of adenovirus and which preferably is introduced into the fiber protein at the level of gene expression (i.e., by introduction of a "nucleic acid sequence that encodes a nonnative amino acid sequence").

Such a nonnative amino acid sequence comprises an amino acid sequence (i.e., has component residues in a particular order) which imparts upon the resultant chimeric protein an ability to bind to and enter cells by means of a novel cell surface binding site (i.e., a "UTV sequence", or Universal Transfer Vector sequence), and/or comprises a sequence incorporated to produce or maintain a certain configuration of the resultant chimeric protein (i.e., a "spacer sequence") between native/nonnative, nonnative/nonnative, or a native/native sequence. Inasmuch as the nonnative amino acid sequence is inserted into or in place of an amino acid sequence, and the manipulation of the amino acid sequence of the chimeric coat protein preferably is made at the nucleic acid level, the amino acid sequence that differs in the chimeric coat protein from the wild-type coat protein (i.e., the UTV sequence and the spacer sequence) preferably can comprise an entirely nonnative amino acid sequence, or a mixture of native and nonnative amino acids).

A "cell surface binding site" encompasses a receptor (which preferably is a protein, carbohydrate, glycoprotein, or proteoglycan) as well as any oppositely charged molecule (i.e., oppositely charged with respect to the chimeric coat protein, which preferably comprises a nonnative amino acid sequence that is positively charged, as described further herein) or other type of molecule with which the chimeric coat protein can interact to bind the cell, and thereby promote cell entry. Examples of potential cell surface binding sites

include, but are not limited to: negatively charged heparin, heparan sulfate, hyaluronic acid, dermatan sulfate, and chondroitin sulfate moieties, for instance, found on glycosaminoglycans, and which further may
5 contain sialic acid, sulfate, and/or phosphate; sialic acid moieties found on mucins, glycoproteins, and gangliosides; major histocompatibility complex I (MHC I) glycoproteins; common carbohydrate components found in membrane glycoproteins, including mannose, N-acetyl-
10 galactosamine, N-acetyl-glucosamine, fucose, galactose, and the like; and phosphate moieties, for instance, on nucleic acids. However, a chimeric coat protein according to the invention, and methods of use thereof, is not limited to any particular mechanism of cellular
15 interaction (i.e., interaction with a particular cell surface binding site) and is not to be so construed.

Furthermore, such a cell surface binding site is "novel" inasmuch as the site is one that previously was inaccessible to interaction with an adenoviral coat
20 protein (i.e., wild-type adenoviral coat protein such as fiber protein), or was accessible only at a very low level, as reflected by the reduced efficiency of entry of a wild-type adenoviral coat protein-containing vector as compared with a vector comprising a chimeric adenovirus
25 coat protein such as fiber protein according to the invention. Moreover, the binding site is novel in that it is present on the majority of, if not all, cells, regardless of their origin. This is in contrast to the cellular binding site with which wild-type adenoviral
30 fiber protein is presumed to interact, which ostensibly is present only on a subset of cells, or is only accessible on a subset of cells, as reflected by the reduced efficiency of entry of a wild-type adenoviral fiber-containing vector.

35 "Efficiency of entry" can be quantitated by several means. In particular, efficiency of entry can be quantitated by introducing a chimeric coat protein into a

vector, preferably a viral vector, and monitoring cell entry (e.g., by vector-mediated delivery to a cell of a gene such as a reporter gene) as a function of multiplicity of infection (MOI)). In this case, a
5 reduced MOI required for cell entry of a vector comprising a chimeric adenoviral coat protein as compared with a vector that is identical except for comprising a wild-type adenoviral coat protein rather than said chimeric adenovirus coat protein, indicates "more
10 efficient" entry.

Similarly, efficiency of entry can be quantitated in terms of the ability of vectors containing chimeric or wild-type coat proteins, or the soluble chimeric or wild-type coat proteins themselves, to bind to cells. In this
15 case, increased binding exhibited for the vector containing a chimeric adenoviral coat protein, or the chimeric coat protein itself, as compared with the identical vector containing a wild-type coat protein instead, or the wild-type coat protein itself, is
20 indicative of an increased efficiency of entry, or "more efficient" entry.

A nonnative amino acid sequence according to the invention preferably is inserted into or in place of an internal coat protein sequence. Alternately, preferably
25 a nonnative amino acid sequence according to the invention is at or near the C-terminus of a protein. In particular, when a coat protein according to the invention is a fiber protein, desirably a nonnative amino acid sequence is at or near the C-terminus of the
30 protein. When a coat protein according to the invention is a penton or hexon protein, preferably a nonnative amino acid sequence is within an exposed loop of the protein, e.g., as described in the following Examples, particularly within a hypervariable region in loop 1
35 and/or loop 2 of the adenovirus hexon protein (Crawford-Miksza et al., J. Virol., 70, 1836-1844 (1996)). Thus, desirably a nonnative amino acid sequence is in a region

of a coat protein that is capable of interacting with and binding a cell.

Furthermore, the method of the invention can be employed to create adenoviral vectors that contain UTV or
5 UTV-like sequences in an extended (or spiked) structure, particularly in hexon and/or penton base protein, so as to result in lengthened hexon and/or penton base proteins, wherein the amino acid insertion projects outward from the protein as it is present in a virion
10 capsid. In particular, such spiked coat proteins can be incorporated into a recombinant adenovirus along with a "short-shafted fiber" (further described herein), wherein the shaft of the fiber has been shortened, and, optionally, the knob of the fiber protein has been
15 replaced with a knob (including the trimerization domain) of another serotype adenoviral vector from which the remainder of the fiber protein derived.

Accordingly, the short-shafted fiber protein preferably can be incorporated into an adenovirus having
20 a chimeric penton base protein that comprises a UTV or UTV-like sequence, or having a "spiked" chimeric penton base protein that furthermore optionally can incorporate a UTV or UTV-like sequence. Also, the short-shafted fiber protein can be incorporated into an adenovirus having a
25 chimeric hexon protein that comprises a UTV or UTV-like sequence, or having a "spiked" chimeric hexon protein that furthermore optionally can incorporate a UTV or UTV-like sequence.

Optimally, the nonnative amino acid sequence is
30 linked to the protein by another nonnative amino acid sequence, i.e., by an intervening spacer sequence. A spacer sequence is a sequence that intervenes between the native protein sequence and the nonnative sequence, between a nonnative sequence and another nonnative
35 sequence, or between a native sequence and another native sequence. A spacer sequence preferably is incorporated into the protein to ensure that the nonnative sequence

comprising the cell surface binding site projects from the three dimensional structure of the chimeric protein (especially the three dimensional structure of the chimeric protein as it exists in nature, i.e., as part of a capsid) in such a fashion so as to be able to interact with and bind to cells. When a spacer sequence is inserted into or replaces an internal coat protein sequence, one or more spacer sequences may be present in the chimeric coat protein.

An intervening spacer sequence can be of any suitable length, preferably from about 3 to about 400 amino acids for a spacer sequence added to derive a "spiked" coat protein (as further described in the following Examples), and preferably from about 3 to about 30 amino acids for any other application described herein. A spacer sequence can comprise any amino acids. Optimally, the spacer sequence does not interfere with the functioning of the coat protein in general, and the functioning of the other nonnative amino acid sequence (i.e., the UTV or UTV-like sequence) in particular.

The nonnative amino acid sequence which is not a spacer sequence, i.e., the UTV sequence, also can be of any suitable length, preferably from about 3 to about 30 amino acids (although, optionally, as for the spacer sequence, the UTV sequence can be longer, e.g., up to about 400 amino acids). These amino acids preferably are any positively charged residues that are capable of binding to negatively charged moieties present on the surface of a eukaryotic cell, and optimally are capable of binding to negatively charged moieties that are present on the surface of the majority of (if not all) eukaryotic cells. In particular, such a negatively charged moiety present on the surface of a eukaryotic cell to which the UTV sequence binds includes the aforementioned "cell surface binding site".

Desirably the nonnative amino acid sequence comprises amino acids selected from the group consisting

of lysine, arginine and histidine. Alternately, these amino acids can be negatively charged residues that are capable of binding to positively charged cell surface binding sites, e.g., desirably the nonnative amino acid sequence comprises amino acids selected from the group consisting of aspartate and glutamate.

Thus, the nonnative amino acid sequence of a coat protein preferably comprises a sequence selected from the group consisting of SEQ ID NO:1 (i.e., Lys Lys Lys Lys Lys Lys Lys Lys), SEQ ID NO:2 (i.e., Arg Arg Arg Arg Arg Arg Arg Arg), and SEQ ID NO:3 (i.e., Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa, wherein "Xaa" comprises Lys or Arg), and wherein 1, 2, 3, 4, or 5 residues of the sequence may be deleted at the C-terminus thereof. When the coat protein is a fiber protein, preferably the protein comprises a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4 (i.e., Gly Ser Asn Lys Glu Ser Phe Val Leu Lys Lys Lys Lys Lys Lys), and SEQ ID NO:5 (i.e., Ala Gly Ser Asn Lys Asn Lys Glu Ser Phe Val Leu Lys Lys Lys Lys Lys Lys), and wherein 1, 2, 3, 4, or 5 residues of the sequence may be deleted at the C-terminus thereof.

Also, sequences that bind that bind to heparin may be involved in binding to a heparin-like receptor (Sawitzky et al., Med. Microbiol. Immunol., 182, 285-92 (1993)). Similarly, so-called "heparin binding sequences" may mediate the interaction of the peptide or protein in which they are contained with other cell surface binding sites, such as with cell surface heparan sulfate proteoglycan (Thompson et al., J. Biol. Chem., 269, 2541-9 (1994)). Thus, preferably the nonnative amino acid sequence (i.e., the UTV sequence) comprises these sequences, as well as additional sequences that are capable of recognizing a negatively charged moiety broadly represented on the surface of eukaryotic cells.

In particular, preferably the nonnative amino acid sequence comprises two basic amino acids (frequently Arg)

located about 20 Å apart, facing in opposite directions of an alpha helix (Margalit et al., J. Biol. Chem., 268, 19228-31 (1993); Ma et al., J. Lipid Res., 35, 2049-2059 (1994)). Other basic amino acids desirably are dispersed
5 between these two residues, facing one side, while nonpolar residues face the other side, forming an amphipathic structure, which optimally comprises the motif XBBXB [SEQ ID NO:49] or XBBBXXB [SEQ ID NO:50], where B is a basic amino acid (e.g., Lys, Arg, etc.), and
10 X is any other amino acid.

Also, preferably the UTV nonnative amino acid sequence comprises: the sequence LIGRKK [SEQ ID NO:51], LIGRK [SEQ ID NO:52] or LIGRR [SEQ ID NO:53], which are common heparin binding motifs present in fibronectin and
15 heat shock proteins (Hansen et al., Biochim. Biophys. Acta, 1252, 135-45 (1995)); insertions of 7 residues of either Lys or Arg, or mixtures of Lys and Arg (Fromm et al., Arch. Biochem. Biophys., 323, 279-87 (1995)); the common basic C-terminal region of IGFBP-3 and IGFBP-5 of
20 about 18 amino acids and which comprises a heparin binding sequence (Booth et al., Growth Regul., 5, 1-17 (1995)); either one or both of the two hyaluronan (HA) binding motifs located within a 35 amino acid region of the C-terminus of the HA receptor RHAMM (Yang et al., J.
25 Cell Biochem., 56, 455-68 (1994)); a synthetic peptide (Ala347-Arg361) modeled after the heparin-binding form of *Staphylococcus aureus* vitronectin comprising heparin-binding consensus sequences (Liang et al., J. Biochem., 116, 457-63 (1994)); any one or more of five heparin
30 binding sites between amino acid 129 and 310 of bovine herpesvirus 1 glycoprotein gIII or any one of four heparin binding sites between amino acids 90 and 275 of pseudorabies virus glycoprotein gIII (Liang et al., Virol., 194, 233-43 (1993)); amino acids 134 to 141 of
35 pseudorabies virus glycoprotein gIII (Sawitzky et al., Med. Microbiol. Immunol., 182, 285-92 (1993); heparin binding regions corresponding to charged residues at

positions 279-282 and 292-304 of human lipoprotein lipase (Ma et al., supra); a synthetic 22 residue peptide, N22W, with a sequence NVSPRRARVTDATETTITISW [SEQ ID NO:54] or residues TETTITIS [SEQ ID NO:55] of this synthetic

5 peptide modeled after fibronectin and which exhibit heparin binding properties (Ingham et al., Arch. Biochem. Biophys., 314, 242-246 (1994)); GVEFVCCP [SEQ ID NO:56] motif present in the ectodomain zinc binding site of the Alzheimer beta-amyloid precursor protein (APP), as well

10 as various other APP-like proteins, which modulates heparin affinity (Bush et al., J. Biol. Chem., 229, 26618-21 (1994)); 8 amino acid residue peptides derived from the cross-region of the laminin A chain (Tashiro et al., Biochem. J., 302, 73-9 (1994)); synthetic peptides

15 based on the heparin binding regions of the serine protease inhibitor antithrombin III including peptides F123-G148 and K121-A134 (Tyler-Cross et al., Protein Sci., 3, 620-7 (1994)); a 14 K N-terminal fragment of APP and a region close to the N-terminus (i.e., residues 96-

20 110) proposed as heparin binding regions (Small et al., J. Neurosci., 14, 2117-27 (1994)); a stretch of 21 amino acids of the heparin binding epidermal growth factor-like growth factor (HB-EGF) characterized by a high content of lysine and arginine residues (Thompson et al., J. Biol. Chem., 269, 2541-9 (1994)); a 17 amino acid region

25 comprising the heparin binding region of thrombospondin and corresponding to a hep 1 synthetic peptide (Murphy-Ullrich et al., J. Biol. Chem., 268, 26784-9 (1993)); a 23 amino acid sequence (Y565-A587) of human von

30 Willebrand factor that binds heparin (Tyler-Cross et al., Arch. Biochem. Biophys., 306, 528-33 (1993)); the fibronectin-derived peptide PRARI [SEQ ID NO:57] (and larger peptides comprising this motif, such as WQPPRARI [SEQ ID NO:58]) that binds heparin (Woods et al., Mol. Biol. Cell., 4, 605-613 (1993); the heparin binding

35 region of platelet factor 4 (Sato et al., Jpn. J. Cancer Res., 84, 485-8 (1993); and the K18K sequence in the

fibroblast growth factor receptor tyrosine kinase transmembrane glycoprotein (Kan et al., Science, 259, 1918-21 (1993)).

Moreover, the UTV sequence can comprise other
5 sequences that are described in the Examples which follow. Thus, preferably the UTV sequence is selected from the group consisting of [SEQ ID NO:1], [SEQ ID NO:2], [SEQ ID NO:3], [SEQ ID NO:4], [SEQ ID NO:5], [SEQ ID NO:20], [SEQ ID NO:22], [SEQ ID NO:24], [SEQ ID
10 NO:26], [SEQ ID NO:28], [SEQ ID NO:30], [SEQ ID NO:32], [SEQ ID NO:34], [SEQ ID NO:36], [SEQ ID NO:38], [SEQ ID NO:40], [SEQ ID NO:42], [SEQ ID NO:46], [SEQ ID NO:48], [SEQ ID NO:49], [SEQ ID NO:50], [SEQ ID NO:51], [SEQ ID NO:52], [SEQ ID NO:53], [SEQ ID NO:54], [SEQ ID
15 NO:55], [SEQ ID NO:56], [SEQ ID NO:57], [SEQ ID NO:58], [SEQ ID NO:73], [SEQ ID NO:74], [SEQ ID NO:76], [SEQ ID NO:78], [SEQ ID NO:90], and [SEQ ID NO:93]. These sequences also can be employed wherein 1, 2, or 3 residues of the sequence are deleted at the C- or N-
20 terminus. Also, inasmuch as a spacer sequence can be any sequence of amino acids that does not interfere with the functioning of the protein, according to the invention, any of the aforementioned UTV sequences also can comprise spacer sequences.

25 It also is preferable that the nonnative amino acid sequence comprise amino acid sequences that are "equivalents" of any of the aforementioned sequences (i.e., are "UTV-like sequences"). An equivalent can be a sequence that carries out the same function (with perhaps
30 minor differences in effectiveness), and yet may differ slightly in terms of its amino acid sequence, or other structural features. In particular, an equivalent sequence is one that comprises one or more conservative amino acid substitutions of the sequence. A "conservative
35 amino acid substitution" is an amino acid substituted by an alternative amino acid of similar charge density, hydrophilicity/hydrophobicity, size, and/or configuration

(e.g., Val for Ile). In comparison, a "nonconservative amino acid substitution" is an amino acid substituted by an alternative amino acid of differing charge density, hydrophilicity/hydrophobicity, size, and/or configuration
5 (e.g., Val for Phe).

Nucleic Acid Encoding a Chimeric Coat Protein

As indicated previously, preferably the nonnative amino acid sequence is introduced at the level of DNA.
10 Accordingly, the invention preferably also provides an isolated and purified nucleic acid encoding a coat protein according to the invention, wherein the nucleic acid sequence that encodes the nonnative amino acid sequence comprises a sequence of SEQ ID NO:6 (i.e., GGA
15 TCC AA), which is located prior to the polyadenylation site. Similarly, the invention preferably provides an isolated and purified nucleic acid comprising a sequence selected from the group consisting of SEQ ID NO:7 (i.e., GGA TCC AAT AAA GAA TCG TTT GTG TTA TGT) and SEQ ID NO:8
20 (i.e., GCC GGA TCC AAC AAG AAT AAA GAA TCG TTT GTG TTA), [SEQ ID NO:19], [SEQ ID NO:21], [SEQ ID NO:23], [SEQ ID NO:25], [SEQ ID NO:27], [SEQ ID NO:29], [SEQ ID NO:31], [SEQ ID NO:33], [SEQ ID NO:35], [SEQ ID NO:37], [SEQ ID NO:39], [SEQ ID NO:41], [SEQ ID NO:45], [SEQ ID NO:47],
25 [SEQ ID NO:72], [SEQ ID NO:75], [SEQ ID NO:77], and [SEQ ID NO:89]. The invention further provides conservatively modified variants of these nucleic acids.

A "conservatively modified variant" is a variation on the nucleic acid sequence that results in a
30 conservative amino acid substitution. In comparison, a "nonconservatively modified variant" is a variation on the nucleic acid sequence that results in a nonconservative amino acid substitution. The means of making such modifications are well known in the art, are
35 described in the Examples which follow, and also can be accomplished by means of commercially available kits and vectors (e.g., New England Biolabs, Inc., Beverly, MA;

Clontech, Palo Alto, CA). Moreover, the means of assessing such substitutions (e.g., in terms of effect on ability to bind and enter cells) are described in the Examples herein.

5 The means of making such a chimeric coat protein, particularly the means of introducing the sequence at the level of DNA, is well known in the art, is illustrated in **Figure 2** for a representative chimeric protein, and is described in the Examples that follow. Briefly, the
10 method comprises introducing a sequence (preferably the sequence of SEQ ID NO:6 or a conservatively modified variant thereof) into the sequence of the coat protein. In one preferred embodiment described in the Examples which follow, the introduction is made prior to any stop
15 codon or polyadenylation signal so as to induce a frameshift mutation into the resultant protein, such that the chimeric protein incorporates additional amino acids.

Generally, this can be accomplished by cloning the fiber sequence into a plasmid or some other vector for
20 ease of manipulation of the sequence. Then, restriction sites flanking the sequence at which the frameshift mutation is to be introduced are identified. A double-stranded synthetic oligonucleotide is created from
25 overlapping synthetic single-stranded sense and antisense oligonucleotides (e.g., from the sense and antisense oligonucleotides, respectively, TAT GGA GGA TCC AAT AAA GAA TCG TTT GTG TTA TGT TTC AAC GTG TTT ATT TTT C [SEQ ID NO:9] and AAT TGA AAA ATA AAC ACG TTG AAA CAT AAC ACA AAC GAT TCT TTA TTG GAT CCT CCA [SEQ ID NO:10], as
30 illustrated in **Figure 4**) such that the double-stranded oligonucleotide incorporates the restriction sites flanking the target sequence. The plasmid or other vector is cleaved with the restriction enzymes, and the oligonucleotide sequence having compatible cohesive ends
35 is ligated in to the plasmid or other vector, to replace the wild-type DNA. Other means of *in vitro* site-directed mutagenesis such as are known to those skilled in the

art, and can be accomplished (in particular, using PCR), for instance, by means of commercially available kits, can be used to introduce the mutated sequence into the coat protein coding sequence.

5 Once the sequence is introduced into the chimeric coat protein, the nucleic acid fragment encoding the sequence can be isolated, e.g., by PCR amplification using 5' and 3' primers. For instance, with respect to a chimeric fiber protein, the fragment can be isolated by
10 PCR using the primer TCCC CCCGGG TCTAGA TTA GGA TCC TTC TTG GGC AAT GTA TGA [SEQ ID NO:11], and the primer CGT GTA TCC ATA TGA CAC AGA [SEQ ID NO:12], as illustrated in **Figure 4**. Use of these primers in this fashion results in an amplified chimeric fiber-containing fragment that
15 is flanked by restriction sites (i.e., in this case *NdeI* and *BamHI* sites) that can be used for convenient subcloning of the fragment. Other means of generating a chimeric coat protein also can be employed.

 Thus, the frameshift mutation can be introduced into
20 any part of a coat protein coding sequence. With respect to SEQ ID NO:6, for instance, this sequence can be placed at the region of the coat protein gene that codes for the C-terminus of the protein (i.e., can be added immediately prior to the TAA stop codon), or can be placed earlier
25 into the coding region, such as between codons coding for Ala (i.e., A) and Gln (i.e., Q) to produce the aforementioned coding sequence of SEQ ID NO:8, which encodes a chimeric protein comprising the sequence of SEQ ID NO:5. Similarly, this approach can be employed to
30 introduce a frameshift even earlier in the coding sequence, e.g., either inserted into or in place of an internal (i.e., native) coat protein sequence.

 Moreover, the double-stranded oligonucleotide can also incorporate a further restriction site that also can
35 be employed in manipulating the sequence. For instance, the sequence of SEQ ID NO:6 introduced in the vector comprises a modified *BamHI* site, i.e., the site is

"modified" in that it adds additional nucleotides onto the palindromic recognition sequence. This sequence also can be synthesized to comprise any other restriction site convenient for DNA manipulations. When incorporated into the coat protein coding sequence, the sequence not only introduces a frame shift mutation, but also can be used to introduce other coding sequences into the coat protein gene. In particular, the coding sequences introduced in this fashion can comprise codons for lysine, arginine and histidine, or codons for aspartate and glutamate, either alone, or in any combination. Furthermore, a new translation stop codon can follow these codons for the amino acids, allowing a chimeric protein to be produced that only incorporates a given number of additional amino acids in the nonnative amino acid sequence. The codons for the amino acids and the translation stop codon can be introduced into the frameshift mutation-inducing novel restriction site incorporated into the coat protein by synthesizing oligonucleotides comprising these sequences that are flanked by the restriction site as previously described (e.g., that comprise 5' and 3' *Bam*HI sites), or by other such means that are known to those skilled in the art.

The size of the DNA used to replace the native receptor binding sequence may be constrained, for example, by impeded folding of the fiber or improper assembly of the penton base/fiber complex. DNA encoding the aforementioned amino acid sequences (e.g., lysine, arginine, histidine, aspartate, glutamate, and the like) is preferred for insertion into the fiber gene sequence in which the native receptor binding sequence has been deleted or otherwise mutated. Moreover, other DNA sequences, such as those that encode amino acids for incorporation into spacer sequences, preferably are used to replace the native coat protein coding sequence.

Vector Comprising a Chimeric Coat Protein

A "vector" according to the invention is a vehicle for gene transfer as that term is understood by those skilled in the art. Four types of vectors encompassed by
5 the invention are plasmids, phages, viruses, and liposomes. Plasmids, phages, and viruses can be transferred to a cell in their nucleic acid form, and liposomes can be employed to transfer nucleic acids. Hence, the vectors that can be employed for gene transfer
10 are referred to herein as "transfer vectors".

Preferably, a vector according to the invention is a virus, especially a virus selected from the group consisting of nonenveloped viruses, i.e., nonenveloped RNA or DNA viruses. Also, a virus can be selected from
15 the group consisting of enveloped viruses, i.e., enveloped RNA or DNA viruses. Such viruses preferably comprise a coat protein. Desirably, the viral coat protein is one that projects outward from the capsid such that it is able to interact with cells. In the case of
20 enveloped RNA or DNA viruses, preferably the coat protein is in fact a lipid envelope glycoprotein (i.e., a so-called spike or peplomer).

In particular, preferably a vector is a nonenveloped virus (i.e., either a RNA or DNA virus) from the family
25 *Hepadnaviridae*, *Parvoviridae*, *Papovaviridae*, *Adenoviridae*, or *Picornaviridae*. A preferred nonenveloped virus according to the invention is a virus of the family *Hepadnaviridae*, especially of the genus *Hepadnavirus*. A virus of the family *Parvoviridae*
30 desirably is of the genus *Parvovirus* (e.g., parvoviruses of mammals and birds) or *Dependovirus* (e.g., adeno-associated viruses (AAVs)). A virus of the family *Papovaviridae* preferably is of the subfamily *Papillomavirinae* (e.g., the papillomaviruses including,
35 but not limited to, human papillomaviruses (HPV) 1-48) or the subfamily *Polyomavirinae* (e.g., the polyomaviruses including, but not limited to, JC, SV40 and BK virus). A

virus of the family *Adenoviridae* desirably is of the genus *Mastadenovirus* (e.g., mammalian adenoviruses) or *Aviadenovirus* (e.g., avian adenoviruses). A virus of the family *Picornaviridae* is preferably a hepatitis A virus (HAV), hepatitis B virus (HBV), or a non-A or non-B hepatitis virus.

Similarly, a vector can be an enveloped virus from the family *Herpesviridae* or *Retroviridae*, or can be a Sindbis virus. A preferred enveloped virus according to the invention is a virus of the family *Herpesviridae*, especially of the subfamily or genus *Alphaherpesvirinae* (e.g., the herpes simplex-like viruses), *Simplexvirus* (e.g., herpes simplex-like viruses), *Varicellavirus* (e.g., varicella and pseudorabies-like viruses), *Betaherpesvirinae* (e.g., the cytomegaloviruses), *Cytomegalovirus* (e.g., the human cytomegaloviruses), *Gammapherpesvirinae* (e.g., the lymphocyte-associated viruses), and *Lymphocryptovirus* (e.g., EB-like viruses).

Another preferred enveloped virus is a RNA virus of the family *Retroviridae* (i.e., preferably is a retrovirus), particularly a virus of the genus or subfamily *Oncovirinae*, *Spumavirinae*, *Spumavirus*, *Lentivirinae*, and *Lentivirus*. A RNA virus of the subfamily *Oncovirinae* is desirably a human T-lymphotropic virus type 1 or 2 (i.e., HTLV-1 or HTLV-2) or bovine leukemia virus (BLV), an avian leukosis-sarcoma virus (e.g., Rous sarcoma virus (RSV), avian myeloblastosis virus (AMV), avian erythroblastosis virus (AEV), Rous-associated virus (RAV)-1 to 50, RAV-0), a mammalian C-type virus (e.g., Moloney murine leukemia virus (MuLV), Harvey murine sarcoma virus (HaMSV), Abelson murine leukemia virus (A-MuLV), AKR-MuLV, feline leukemia virus (FeLV), simian sarcoma virus, reticuloendotheliosis virus (REV), spleen necrosis virus (SNV)), a B-type virus (e.g., mouse mammary tumor virus (MMTV)), or a D-type virus (e.g., Mason-Pfizer monkey virus (MPMV), "SAIDS" viruses). A RNA virus of the subfamily *Lentivirus* is

desirably a human immunodeficiency virus type 1 or 2 (i.e., HIV-1 or HIV-2, wherein HIV-1 was formerly called lymphadenopathy associated virus 3 (HTLV-III) and acquired immune deficiency syndrome (AIDS)-related virus (ARV)), or another virus related to HIV-1 or HIV-2 that has been identified and associated with AIDS or AIDS-like disease. The acronym "HIV" or terms "AIDS virus" or "human immunodeficiency virus" are used herein to refer to these HIV viruses, and HIV-related and -associated viruses, generically. Moreover, a RNA virus of the subfamily *Lentivirus* preferably is a Visna/maedi virus (e.g., such as infect sheep), a feline immunodeficiency virus (FIV), bovine lentivirus, simian immunodeficiency virus (SIV), an equine infectious anemia virus (EIAV), or a caprine arthritis-encephalitis virus (CAEV).

An especially preferred vector according to the invention is an adenoviral vector (i.e., a viral vector of the family *Adenoviridae*, optimally of the genus *Mastadenovirus*). Desirably such a vector is an Ad2 or Ad5 vector, although other serotype adenoviral vectors can be employed. The adenoviral vector employed for gene transfer can be wild-type (i.e., replication competent). Alternately, the adenoviral vector can comprise genetic material with at least one modification therein, which can render the virus replication deficient. The modification to the adenoviral genome can include, but is not limited to, addition of a DNA segment, rearrangement of a DNA segment, deletion of a DNA segment, replacement of a DNA segment, or introduction of a DNA lesion. A DNA segment can be as small as one nucleotide and as large as 36 kilobase pairs (i.e., the approximate size of the adenoviral genome) or, alternately, can equal the maximum amount which can be packaged into an adenoviral virion (i.e., about 38 kb). Preferred modifications to the adenoviral genome include modifications in the E1, E2, E3 or E4 region. Similarly, an adenoviral vector can be a cointegrate, i.e., a ligation of adenoviral sequences,

with other sequences, such as other virus or plasmid sequences.

In terms of a viral vector (e.g., particularly a replication deficient adenoviral vector), such a vector
5 can comprise either complete capsids (i.e., including a viral genome such as an adenoviral genome) or empty capsids (i.e., in which a viral genome is lacking, or is degraded, e.g., by physical or chemical means). Along the same lines, since methods are available for
10 transferring viruses, plasmids, and phages in the form of their nucleic acid sequences (i.e., RNA or DNA), a vector (i.e., a transfer vector) similarly can comprise RNA or DNA, in the absence of any associated protein such as capsid protein, and in the absence of any envelope lipid.
15 Similarly, since liposomes effect cell entry by fusing with cell membranes, a transfer vector can comprise liposomes (e.g., such as are commercially available, for instance, from Life Technologies, Bethesda, MD), with constitutive nucleic acids encoding the coat protein.
20 Thus, according to the invention whereas a vector "comprises" a chimeric adenoviral coat protein, a transfer vector "encodes" a chimeric adenoviral coat protein; liposome transfer vectors in particular "encode" in the sense that they contain nucleic acids which, in
25 fact, encode the protein.

A vector according to the invention can comprise additional sequences and mutations, e.g., some within the coat protein itself. For instance, a vector according to the invention further preferably comprises a nucleic acid
30 comprising a passenger gene.

A "nucleic acid" is a polynucleotide (DNA or RNA). A "gene" is any nucleic acid sequence coding for a protein or a nascent RNA molecule. A "passenger gene" is any gene which is not typically present in and is
35 subcloned into a vector (e.g., a transfer vector) according to the present invention, and which upon introduction into a host cell is accompanied by a

discernible change in the intracellular environment (e.g., by an increased level of deoxyribonucleic acid (DNA), ribonucleic acid (RNA), peptide or protein, or by an altered rate of production or degradation thereof). A

5 "gene product" is either an as yet untranslated RNA molecule transcribed from a given gene or coding sequence (e.g., mRNA or antisense RNA) or the polypeptide chain (i.e., protein or peptide) translated from the mRNA molecule transcribed from the given gene or coding

10 sequence. Whereas a gene comprises coding sequences plus any non-coding sequences, a "coding sequence" does not include any non-coding (e.g., regulatory) DNA. A gene or coding sequence is "recombinant" if the sequence of bases along the molecule has been altered from the sequence in

15 which the gene or coding sequence is typically found in nature, or if the sequence of bases is not typically found in nature. According to this invention, a gene or coding sequence can be wholly or partially synthetically made, can comprise genomic or complementary DNA (cDNA)

20 sequences, and can be provided in the form of either DNA or RNA.

Non-coding sequences or regulatory sequences include promoter sequences. A "promoter" is a DNA sequence that directs the binding of RNA polymerase and thereby

25 promotes RNA synthesis. "Enhancers" are *cis*-acting elements of DNA that stimulate or inhibit transcription of adjacent genes. An enhancer that inhibits transcription is also termed a "silencer". Enhancers differ from DNA-binding sites for sequence-specific DNA

30 binding proteins found only in the promoter (which are also termed "promoter elements") in that enhancers can function in either orientation, and over distances of up to several kilobase pairs, even from a position downstream of a transcribed region. According to the

35 invention, a coding sequence is "operably linked" to a promoter (e.g., when both the coding sequence and the promoter constitute a passenger gene) when the promoter

is capable of directing transcription of that coding sequence.

Accordingly, a "passenger gene" can be any gene, and desirably is either a therapeutic gene or a reporter gene. Preferably a passenger gene is capable of being expressed in a cell in which the vector has been internalized. For instance, the passenger gene can comprise a reporter gene, or a nucleic acid sequence which encodes a protein that can in some fashion be detected in a cell. The passenger gene also can comprise a therapeutic gene, for instance, a therapeutic gene which exerts its effect at the level of RNA or protein. For instance, a protein encoded by a transferred inherited disease, such as, e.g., the cystic fibrosis transmembrane conductance regulator cDNA for the treatment of cystic fibrosis. The protein encoded by the therapeutic gene may exert its therapeutic effect by resulting in cell killing. For instance, expression of the gene in itself may lead to cell killing, as with expression of the diphtheria toxin A gene, or the expression of the gene may render cells selectively sensitive to the killing action of certain drugs, e.g., expression of the HSV thymidine kinase gene renders cells sensitive to antiviral compounds including acyclovir, gancyclovir and FIAU (1-(2-deoxy-2-fluoro- β -D-arabinofuranosil)-5-iodouracil).

Moreover, the therapeutic gene can exert its effect at the level of RNA, for instance, by encoding an antisense message or ribozyme, a protein which affects splicing or 3' processing (e.g., polyadenylation), or can encode a protein which acts by affecting the level of gene expression of another gene within the cell (i.e., where expression is broadly considered to include all steps from initiation of transcription through production of a processed protein), perhaps, among other things, by mediating an altered rate of mRNA accumulation, an

alteration of mRNA transport, and/or a change in post-transcriptional regulation. Accordingly, the use of the term "therapeutic gene" is intended to encompass these and any other embodiments of that which is more commonly referred to as gene therapy and is known to those of skill in the art. Similarly, the recombinant adenovirus can be used for gene therapy or to study the effects of expression of the gene in a given cell or tissue *in vitro* or *in vivo*.

10 The recombinant adenovirus comprising a chimeric coat protein such as a fiber protein and the recombinant adenovirus that additionally comprises a passenger gene or genes capable of being expressed in a particular cell can be generated by use of a transfer vector, preferably
15 a viral or plasmid transfer vector, in accordance with the present invention. Such a transfer vector preferably comprises a chimeric adenoviral coat protein gene sequence as previously described. The chimeric coat protein gene sequence comprises a nonnative sequence in
20 place of the native sequence, which has been deleted, or in addition to the native sequence.

 A recombinant chimeric coat protein gene sequence (such as a fiber gene sequence) can be moved from an adenoviral transfer vector into baculovirus or a suitable
25 prokaryotic or eukaryotic expression vector for expression and evaluation of receptor or protein specificity and avidity, trimerization potential, penton base binding, and other biochemical characteristics.

 Accordingly, the present invention also provides
30 recombinant baculoviral and prokaryotic and eukaryotic expression vectors comprising a chimeric adenoviral coat protein gene sequence (preferably a fiber gene sequence), which also are "transfer vectors" as defined herein. The chimeric coat protein gene sequence (e.g., fiber gene
35 sequence) includes a nonnative sequence in addition to or in place of a native amino acid sequence, and which enables the resultant chimeric coat protein (e.g., fiber

protein) to bind to a binding site other than a binding site bound by the native sequence. By moving the chimeric gene from an adenoviral vector to baculovirus or a prokaryotic or eukaryotic expression vector, high protein expression is achievable (approximately 5-50% of the total protein being the chimeric fiber).

A vector according to the invention further can comprise, either within, in place of, or outside of the coding sequence of a coat protein additional sequences that impact upon the ability of a coat protein such as fiber protein to trimerize, or comprise a protease recognition sequence. A sequence that impacts upon the ability to trimerize is one or more sequences that enable trimerization of a chimeric coat protein that is a fiber protein. A sequence that comprises a protease recognition sequence is a sequence that can be cleaved by a protease, thereby effecting removal of the chimeric coat protein (or a portion thereof) and attachment of the recombinant adenovirus to a cell by means of another coat protein. When employed with a coat protein that is a fiber protein, the protease recognition or receptor does not affect fiber trimerization, preferably the fiber specificity of the fiber protein. For instance, in one embodiment of the present invention, deletion of the fiber protein, or a portion thereof, is deleted in by means of a protease recognition sequence, and then the novel cell surface binding site is incorporated into either the penton base or hexon coat protein, preferably with use of a spacer sequence as previously described.

In terms of the production of vectors and transfer vectors according to the invention, transfer vectors are constructed using standard molecular and genetic techniques such as are known to those skilled in the art. Vectors (e.g., virions or virus particles) are produced using viral vectors. For instance, a viral vector comprising a chimeric coat protein according to the invention can be constructed by providing to a cell that

does not comprise any E4 complementing sequences: (1) a linear vector comprising the chimeric fiber and the wild-type E4 gene, and (2) a linear vector that is E4', as illustrated in Figure 3. As described in the Examples which follow, this methodology results in recombination between the sequences, generating a vector that comprises a portion of the initial E4' vector and a portion of the E4' vector, particularly the region comprising the chimeric fiber sequences.

Similarly, the fiber chimera-containing particles are produced in standard cell lines, e.g., those currently used for adenoviral vectors. Following production and purification, the particles in which fiber is to be deleted are rendered fiberless through digestion of the particles with a sequence-specific protease, which cleaves the fiber proteins and releases them from the viral particles to generate fiberless particles. For example, thrombin recognizes and cleaves at known amino acid sequences that can be incorporated into the vector (Stenflo et al., J. Biol. Chem., 257, 12280-12290 (1982)). Similarly, deletion mutants lacking the fiber gene can be employed in vector construction, e.g., H2d1802, H2d1807, and H2d11021 (Falgout et al., J. Virol., 62, 622-625 (1988)). These fiberless particles have been shown to be stable and capable of binding and infecting cells (Falgout et al., supra). These resultant particles then can be targeted to specific tissues via the penton base or other coat protein, preferably such other coat protein that comprises one or more nonnative amino acid sequences according to the invention.

Alternately, recombinant adenovirus comprising chimeric fiber protein having further modifications can be produced by the removal of the native knob region, which comprises receptor-binding and trimerization domains, of the fiber protein and its replacement with a nonnative trimerization domain (Peteranderl et al., Biochemistry, 31, 12272-12276 (1992)) and a nonnative

amino acid sequence according to the invention. A recombinant adenovirus comprising a chimeric fiber protein also can be produced by point mutation in the knob region and the isolation of clones that are capable of trimerization. In either case, and also with respect to the removal and replacement of the native receptor-specific binding sequence as described above, new protein binding domains can be added onto the C-terminus of the fiber protein or into exposed loops of the fiber protein by inserting one or more copies of the nucleic acid sequence encoding the nonnative amino acid sequence into the appropriate position. Preferably, such a fiber protein is able to trimerize, so that it is able to bind to penton base protein.

The method described above for generating chimeric fiber protein also can be used to make other chimeric coat proteins, e.g., chimeric hexon or penton protein.

Illustrative Uses

The present invention provides a chimeric protein that is able to bind to cells and mediate entry into cells with high efficiency, as well as vectors and transfer vectors comprising same. The chimeric coat protein itself has multiple uses, e.g., as a tool for studies *in vitro* of adenovirus binding to cells (e.g., by Scatchard analysis as shown previously by Wickham et al. (1993), supra), to block binding of adenovirus to receptors *in vitro* (e.g., by using antibodies, peptides, and enzymes, as described in the Examples), and to protect against adenoviral infection *in vivo* by competing for binding to the binding site by which adenovirus effects cell entry.

A vector comprising a chimeric coat protein also can be used in strain generation and as a means of making new vectors. For instance, the nonnative amino acid sequence can bind to nucleic acids, and can be introduced intracellularly as a means of generating new vectors via

recombination. Similarly, a vector can be used in gene therapy. For instance, a vector of the present invention can be used to treat any one of a number of diseases by delivering to targeted cells corrective DNA, i.e., DNA
5 encoding a function that is either absent or impaired, or a discrete killing agent, e.g., DNA encoding a cytotoxin that, for example, is active only intracellularly. Diseases that are candidates for such treatment include, for example, cancer, e.g., melanoma, glioma or lung
10 cancers; genetic disorders, e.g., cystic fibrosis, hemophilia or muscular dystrophy; pathogenic infections, e.g., human immunodeficiency virus, tuberculosis or hepatitis; heart disease, e.g., preventing restenosis following angioplasty or promoting angiogenesis to
15 reperfuse necrotic tissue; and autoimmune disorders, e.g., Crohn's disease, colitis or rheumatoid arthritis.

In particular, gene therapy can be carried out in the treatment of diseases, disorders, or conditions associated with different tissues that ostensibly lack
20 high levels of the receptor to which wild-type adenovirus fiber protein binds, and thus for which current adenoviral-mediated approaches to gene therapy are less than optimal (e.g., for delivery to monocyte/macrophages, fibroblasts, neuronal, smooth muscle, and epithelial
25 cells). Tissues comprised of these cells (and diseases, disorders, or conditions associated therewith) include, but are not limited to: endothelia (e.g., angiogenesis, restenosis, inflammation, and tumors); neuronal tissue (e.g., tumors and Alzheimer's disease); epithelium (e.g.,
30 disorders of the skin, cornea, intestine, and lung); hematopoietic cells (e.g., human immunodeficiency virus (HIV-1, HIV-2), cancers, and anemias); smooth muscle (e.g., restenosis); and fibroblasts (e.g., inflammation).

Moreover, instead of transferring a therapeutic gene,
35 a reporter gene, or some type of marker gene can be transferred instead using the vectors (particularly the adenoviral vectors) of the invention. Marker genes and

reporter genes are of use, for instance, in cell differentiation and cell fate studies, as well as potentially for diagnostic purposes. Moreover, a standard reporter gene such as a β -galactosidase reporter gene, a gene encoding green fluorescent protein (GFP), or a β -glucuronidase gene can be used *in vivo*, e.g., as a means of assay in a living host, or, for instance, as a means of targeted cell ablation (see, e.g., Minden et al., Biotechniques, 20, 122-129 (1996); Youvan, Science, 268, 264 (1995); U.S. Patent 5,432,081; Deonarain et al., Br. J. Cancer, 70, 786-794 (1994)).

Similarly, it may be desirable to transfer a gene to use a host essentially as a means of production *in vivo* of a particular protein. Along these lines, transgenic animals have been employed, for instance for the production of recombinant polypeptides in the milk of transgenic bovine species (e.g., PCT International Application WO 93/25567). Other "non-therapeutic" reasons for gene transfer include the study of human diseases using an animal model (e.g., use of transgenic mice and other transgenic animals including p53 tumor suppressor gene knockouts for tumorigenic studies, use of a transgenic model for impaired glucose tolerance and human Alzheimer's amyloid precursor protein models for the study of glucose metabolism and pathogenesis of Alzheimer's disease, respectively, etc.)

These aforementioned illustrative uses are by no means comprehensive, and it is intended that the present invention encompasses such further uses which flow from, but are not explicitly recited, in the disclosure herein. Similarly, there are numerous advantages associated with the use of the various aspects of the present invention.

For instance, use of a universal targeting vector according to the invention is advantageous inasmuch as:

- (1) the vector can potentially be used for all cells and tissues;
- (2) only one vector is required for use in all cell lines, there is no need for co-transfecting an

independent vector; (3) the vector is capable of effecting gene delivery with an efficiency that is increased over that observed for vectors comprising wild-type fiber protein; (4) the vector, unlike prior vectors, does not target specific cells, but instead increases transduction efficiency in what appears to be a global fashion; (5) the vector is capable of mediating gene transfer when employed at a reduced dose (i.e., multiplicity of infection (MOI)) as compared with vector comprising wild-type fiber protein, and thus likely reduces the dosage-related drawbacks that accompany currently available adenoviral vectors; and (6) the vector can be propagated and maintained using currently available cell lines.

The ability of a universal targeting vector such as a universal targeting adenovirus vector to potentially bind to and enter all or most tissues has several advantages. These advantages include increased gene delivery efficiency to multiple tissues, the availability of a single vector capable of delivering genes to all tissues, and simplified production of necessary components for gene delivery. Moreover, such a universal targeting vector comprises a potential to deliver exogenous DNA into cells by "piggy backing" the DNA on the vector by means of a protein/DNA interaction.

Further potential advantages of such a universal targeting vector include a substantially increased efficiency of delivery (e.g., increased by 10- to 100-fold) into cells expressing low levels of fiber receptor to which wild-type fiber protein binds, as well as increased efficiency into cells or tissues expressing fiber receptor to which wild-type fiber binds. Moreover the reduced dosage at which the vectors are employed should result in a decrease in adenovirus-associated inflammation, the humoral response to adenovirus, and the cytotoxic T-lymphocyte response to adenovirus.

Furthermore, the vector is advantageous in that it can be isolated and purified by conventional means. Since changes in the vector are made at the genome level, there are no cumbersome and costly post-production
5 modifications required, as are associated with other vectors (see, e.g., Cotten et al., supra; Wagner et al., supra). Similarly, special receptor-expressing cells lines are not required. A UTV vector can be propagated to similar titers as a wild-type vector lacking the fiber
10 modification.

Means of Administration

The vectors and transfer vectors of the present invention can be employed to contact cells either *in*
15 *vitro* or *in vivo*. According to the invention "contacting" comprises any means by which a vector is introduced intracellularly; the method is not dependent on any particular means of introduction and is not to be so construed. Means of introduction are well known to
20 those skilled in the art, and also are exemplified herein.

Accordingly, introduction can be effected, for instance, either *in vitro* (e.g., in an *ex vivo* type method of gene therapy or in tissue culture studies) or
25 *in vivo* by electroporation, transformation, transduction, conjugation or triparental mating, (co-)transfection, (co-)infection, membrane fusion with cationic lipids, high velocity bombardment with DNA-coated microprojectiles, incubation with calcium phosphate-DNA
30 precipitate, direct microinjection into single cells, and the like. Similarly, the vectors can be introduced by means of cationic lipids, e.g., liposomes. Such liposomes are commercially available (e.g., Lipofectin®, Lipofectamine™, and the like, supplied by Life
35 Technologies, Gibco BRL, Gaithersburg, MD). Moreover, liposomes having increased transfer capacity and/or reduced toxicity *in vivo* (see, e.g., PCT patent

application WO 95/21259) can be employed in the present invention. Other methods also are available and are known to those skilled in the art.

According to the invention, a "host" (and thus a "cell" from a host) encompasses any host into which a vector of the invention can be introduced, and thus encompasses an animal, including, but not limited to, an amphibian, bird, fish, insect, reptile, or mammal. Optimally a host is a mammal, for instance, rodent, primate (such as chimpanzee, monkey, ape, gorilla, orangutan, or gibbon), feline, canine, ungulate (such as ruminant or swine), as well as, in particular, human.

Inasmuch as a universal targeting vector (such as enters all cells, a cell can be any cell into which such a vector can enter. In particular, a universal targeting vector can be employed for gene transfer to a cell that expresses low or undetectable levels of fiber receptor, including, but not limited to, an endothelial, smooth muscle, neuronal, hematopoietic, or fibroblast cell.

One skilled in the art will appreciate that suitable methods of administering a vector (particularly an adenoviral vector) of the present invention to an animal for purposes of gene therapy (see, for example, Rosenfeld et al., Science, 252, 431-434 (1991); Jaffe et al., Clin. Res., 39(2), 302A (1991); Berkner, BioTechniques, 5, 616-629 (1988)), chemotherapy, and vaccination are available, and, although more than one route can be used for administration, a particular route can provide a more immediate and more effective reaction than another route. Pharmaceutically acceptable excipients also are well-known to those who are skilled in the art, and are readily available. The choice of excipient used to administer the recombinant vector. Accordingly, there is a wide variety of suitable formulations for use in the context of the present invention. The following methods

and excipients are merely exemplary and are in no way limiting.

Formulations suitable for oral administration can consist of (a) liquid solutions, such as an effective amount of the compound dissolved in diluents, such as water, saline, or orange juice; (b) capsules, sachets or tablets, each containing a predetermined amount of the active ingredient, as solids or granules; (c) suspensions in an appropriate liquid; and (d) suitable emulsions. Tablet forms can include one or more of lactose, mannitol, corn starch, potato starch, microcrystalline cellulose, acacia, gelatin, colloidal silicon dioxide, croscarmellose sodium, talc, magnesium stearate, stearic acid, and other excipients, diluents, buffering agents, moistening agents, preservatives, flavoring agents, and pharmacologically compatible excipients. Lozenge forms can comprise the active ingredient in a flavor, usually sucrose and acacia or tragacanth, as well as pastilles comprising the active ingredient in an inert base, such as gelatin and glycerin, or sucrose and acacia, emulsions, gels, and the like containing, in addition to the active ingredient, a vector or transfer vector of the present invention, alone or in combination with other suitable components, can be made into aerosol formulations to be administered via inhalation. These aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and the like. They may also be formulated as pharmaceuticals for non-pressured preparations such as in a nebulizer or an atomizer.

Formulations suitable for parenteral administration include aqueous and non-aqueous, isotonic sterile injection solutions, which can contain anti-oxidants, buffers, bacteriostats, and solutes that render the formulation isotonic with the blood of the intended

recipient, and aqueous and non-aqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. The formulations can be presented in unit-dose or multi-dose sealed containers, such as ampules and vials, and can be stored in a freeze-dried (lyophilized) condition requiring only the addition of the sterile liquid excipient, for example, water, for injections, immediately prior to use. Extemporaneous injection solutions and suspensions can be prepared from sterile powders, granules, and tablets of the kind previously described.

Additionally, a vector or transfer vector of the present invention can be made into suppositories by mixing with a variety of bases such as emulsifying bases or water-soluble bases.

Formulations suitable for vaginal administration can be presented as pessaries, tampons, creams, gels, pastes, foams, or spray formulas containing, in addition to the active ingredient, such carriers as are known in the art to be appropriate.

The dose administered to an animal, particularly a human, in the context of the present invention will vary with the gene of interest, the composition employed, the method of administration, and the particular site and organism being treated. However, the dose should be sufficient to effect a therapeutic response.

As previously indicated, a vector or a transfer vector of the present invention also has utility *in vitro*. Such a vector can be used as a research tool in the study of adenoviral attachment and infection of cells and in a method of assaying binding site-ligand interaction. Similarly, the recombinant coat protein comprising a nonnative amino acid sequence in addition to or in place of a native receptor binding sequence can be used in receptor-ligand assays and as adhesion proteins *in vitro* or *in vivo*, for example.

The following examples further illustrate the present invention and, of course, should not be construed as in any way limiting its scope.

5

Example 1

This example describes an investigation of the levels of adenovirus receptor in different cells, as determined by the ability of wild-type adenovirus to bind to the cells.

10

For these experiments, the ability of adenovirus comprising wild-type fiber to bind to cells derived from various tissues was assessed. Adenovirus particles of an Ad5 strain were labeled with [³H]-thymidine as previously described (see, e.g., Wickham et al., Cell, 73, 309-319 (1993)). Subsatürating levels of thymidine-labeled adenovirus were added to 200 µl of 10⁶ cells preincubated about 30 to 60 minutes with or without 20 µg/ml of soluble fiber protein. The cells were incubated with the virus for 1 hour at 4°C and then washed 3 times with cold phosphate buffered saline (PBS). The remaining cell-associated counts were measured in a scintillation counter. Specific binding was measured by subtracting the cell-associated counts (i.e., counts per minute (cpm)) in the presence of fiber from the cell-associated counts in the absence of fiber. Binding in the presence of fiber was never more than 2% of the total input of radioactive virus particles. Results were obtained as the average of triplicate measurements.

20

25

As illustrated in **Figure 1**, a substantial number of the cells derived from different tissues expressed little or no fiber receptor, as indicated by a relative inability of wild-type adenovirus to bound to these cells. Cells of epithelial origin (i.e., "receptor-plus" cells including Chang, HeLa, and A549 cells) bound high levels of adenovirus. In comparison, non-epithelial cells (i.e., "receptor-minus" cells such as monocyte/macrophages, fibroblasts, neuronal, smooth

30

35

muscle, and epithelial cells) exhibited about 10-fold or more reductions in virus binding as compared to epithelial-like cells.

These results confirm the previously unrecognized relative inability of adenovirus to bind to and hence enter receptor-minus non-epithelial cells, as compared with receptor-plus epithelial cells. Presumably this inability is due to the low representation of receptors for wild-type adenoviral fiber protein on these cells.

Example 2

This example describes the construction of an adenoviral vector comprising a chimeric coat protein, particularly a chimeric adenoviral fiber protein.

To overcome the transduction limitation imposed by the presence of only a limited number of fiber receptors on clinically relevant tissues such as non-epithelial tissue, a modified adenovirus vector was constructed as depicted in Figures 2A and 2B to derive a vector that is referred to herein as a "universal transfer vector", or UTV. In particular, a frameshift mutation was created in a gene encoding an adenoviral coat protein, in this case, in the fiber gene. In wild-type adenovirus, the unmodified fiber gene contains a nested translational stop signal (TAA) and transcriptional polyadenylation signal (AATAAA). The polyadenylation signal directs the addition of a polyA tail onto the 3' end of the transcript. The polyA tail typically comprises anywhere from about 20 to about 200 nucleotides. Following transcription and exit from the nucleus, the TAA stop signal directs termination of translation by the ribosome.

In comparison, the modified fiber gene of a UTV vector lacks an in-frame translational "stop" signal. Following normal transcription and addition of the polyA extension onto the mRNA, in the absence of the stop codon, the ribosome continues translation of the

transcript into the polyA region. Inasmuch as the codon AAA codes for the amino acid lysine, the resultant chimeric fiber gene translation product produced by a UTV contains an addition of a string of polylysine residues at the C-terminus, i.e., Lys Lys Lys Lys Lys Lys Lys Lys [SEQ ID NO:1]. It is possible that a cellular process acts to limit the length of the polylysine string, since the polylysine residues typically comprise from about 3 to about 30 residues in the chimeric fiber protein.

Whatever the case, however, the polylysine protein modification, as well as further modifications described herein, allows the UTV to efficiently attach to cells lacking high levels of the receptor for wild-type adenoviral fiber protein (i.e., receptor-minus cells).

In terms of vector construction and characterization, standard molecular and genetic techniques, such as the generation of strains, plasmids, and viruses, gel electrophoresis, DNA manipulations including plasmid isolation, DNA cloning and sequencing, Western blot assays, and the like, were performed such as are known to those skilled in the art, and as are described in detail in standard laboratory manuals (e.g., Maniatis et al., Molecular Cloning: A Laboratory Manual, 2nd ed. (Cold Spring Harbor, NY, 1992); Ausubel et al., Current Protocols in Molecular Biology, (1987)).

Restriction enzymes and other enzymes used for molecular manipulations were purchased from commercial sources (e.g., Boehringer Mannheim, Inc., Indianapolis, Indiana; New England Biolabs, Beverly, Massachusetts; Bethesda Research Laboratories, Bethesda, Maryland), and were used according to the recommendations of the manufacturer. Cells employed for experiments (e.g., cells of the transformed human embryonic kidney cell line 293 (i.e., CRL 1573 cells) and other cells supplied by American Type Culture Collection) were cultured and maintained using standard sterile culture reagents, media and techniques,

as previously described (Erzerum et al., Nucleic Acids Research, 21, 1607-1612 (1993)).

Accordingly, the frameshift mutation of the fiber stop codon was created by introducing a modified *Bam*HI site (i.e., GGATCCAA [SEQ ID NO:6]) into an adenoviral transfer vector. This was done as illustrated in Figure 3 by starting with the transfer plasmid pAd NS 83-100 (which also is known as p193NS 83-100 or pNS 83-100). pAd NS 83-100 was constructed by cloning the Ad5 *Nde*I to *Sal*I fragment, which spans the 83-100 map unit region of the Ad5 genome containing the fiber gene, into the plasmid pNEB193 (New England Biolabs, Beverly, MA).

The *Nde*I-*Mun*I fragment of pAd NS 83-100 was replaced with a synthetic oligonucleotide comprising a *Bam*HI site, which was flanked by a 5' *Nde*I site and a 3' *Mun*I site to facilitate cloning. The double-stranded synthetic oligonucleotide fragment was created from overlapping synthetic single-stranded sense and antisense oligonucleotides, i.e., respectively, the sense primer TAT GGA GGA TCC AAT AAA GAA TCG TTT GTG TTA TGT TTC AAC GTG TTT ATT TTT C [SEQ ID NO:9], and the antisense primer AAT TGA AAA ATA AAC ACG TTG AAA CAT AAC ACA AAC GAT TCT TTA TTG GAT CCT CCA [SEQ ID NO:10], as illustrated in Figures 4A and 4B, respectively. The ends of the overlapping oligomers were made to have overhangs compatible for direct cloning into the *Nde*I and *Mun*I sites.

The resultant transfer plasmid, pAd NS 83-100 (F⁻) (which also is known as p193NS (Δ F) or pNS (Δ F)), lacks all but the first 50 base pairs of the coding sequence for the fiber gene (i.e., is "fiber-minus"). The vector furthermore contains the entire adenovirus E4 coding sequence. The plasmid retains the AATAAA polyadenylation signal included in the synthetic *Nde*I/*Mun*I oligonucleotide, and also incorporates the new *Bam*HI restriction site.

The mutated fiber gene was incorporated into the fiber-minus pAd NS 83-100 plasmid using synthetic sense and antisense oligonucleotide primers to amplify the fiber gene with use of the polymerase chain reaction (PCR) while incorporating a modified *Bam*HI site following the last codon of the fiber gene to create the mutant fiber gene. This incorporated modified *Bam*HI site also serves to code for the amino acids glycine and serine, resulting in a chimeric nucleic acid sequence of GGA TCC AAT AAA GAA TCG TTT GTG TTA TGT [SEQ ID NO:7]. The modified fiber gene thus codes for an extension to the resultant chimeric fiber protein of Gly Ser Asn Lys Glu Ser Phe Val Leu Lys Lys Lys [SEQ ID NO:4], wherein the length of the polylysine string can vary. The synthetic oligonucleotides employed for fiber amplification were the primer TCCC CCCGGG TCTAGA TTA GGA TCC TTC TTG GGC AAT GTA TGA [SEQ ID NO:11], and the primer CGT GTA TCC ATA TGA CAC AGA [SEQ ID NO:12], as illustrated in Figures 4C and 4D, respectively.

The amplified gene product was then cut with the restriction enzymes *Nde*I and *Bam*HI, and was cloned into the *Nde*I/*Bam*HI sites of the fiber-minus plasmid pAd NS 83-100 to create the transfer vector pAd NS 83-100 UTV (which also is known as p193NS (F5*), p193 (F5*), or pNS (F5*)). The entire *Nde*I to *Sal*I adenovirus sequence of pAd NS 83-100 UTV was cloned into the fiber-minus plasmid pAd BS 59-100 to create pAd BS 59-100 UTV (which also is known as p193NS (F5*), p193 (F5*), or pNS (F5*)).

The UTV adenovirus vector was created through homologous recombination in 293 cells. Namely, the E4' pAd BS 59-100 UTV transfer vector was linearized with *Sal*I, and was transfected into 293 cells that were previously infected with the adenovirus vector, A2F. The A2F vector was derived from a GV10 vector. The Ad5-based vector GV10 contains the *lacZ* gene under the control of the Rous sarcoma virus promoter (i.e., comprises RSV *lacZ*). The insertion of the reporter gene in GV10 is

made within the E1 region (i.e., the vector is E1⁻). The GV10 vector also contains a deletion of the E3 region, but is E4⁺. In comparison with GV10 (i.e., RSV lacZ E1⁻ E3⁻ E4⁺), A2F further comprises a deletion of the essential E4 adenovirus genes, but is E3⁺ (i.e., RSV lacZ E1⁻ E3⁺ E4⁻).

The 293 cells contain an E1 complementing sequence, but do not contain an E4 complementing sequence. The lack of an E4 complementing sequence prevents replication of the E4⁺ A2F vector in the 293 cell line. However, upon co-introduction of A2F virus and pAd BS 59-100 UTV in 293 cells, homologous recombination takes place between the UTV transfer vector and the A2F adenoviral genome, producing an E3⁺ E4⁺ adenovirus genome comprising a chimeric fiber protein, which is capable of replication in 293 cells. This particular resultant UTV vector was designated GV10 UTV.

The GV10 UTV vector was isolated using standard plaque isolation techniques with 293 cells. Following three successive rounds of plaque-purification, the GV10 UTV vector contained the fiber mutation and was free of any contamination by the E4⁺ A2F vector. The presence of the chimeric fiber sequences in the GV10 UTV vector was confirmed by sequencing the fiber mRNA using reverse transcriptase-polymerase chain reaction (RT-PCR), which validated the presence of a polyadenine tail in the chimeric fiber mRNA.

Similarly, the production of a chimeric fiber protein by the vector was confirmed by Western blot. To accomplish this, 293 cells were infected at a multiplicity of infection (MOI) of 5 with either GV10 comprising wild-type adenoviral fiber protein or with GV10 UTV comprising chimeric fiber protein. At two days post-infection, the cells were washed and then lysed in PBS by three freeze-thaw cycles. The lysates were cleared by centrifugation and loaded onto a 10% sodium dodecyl sulfate/polyacrylamide gel. Following

electrophoresis, the proteins were transferred onto nitrocellulose and detected by chemiluminescence using a polyclonal antibody to fiber. The Western blot is depicted in **Figure 5**. As can be seen from this figure, the migration of the proteins indicates that the chimeric UTV fiber is about 1.5 to about 2.0 kilodaltons larger than the unmodified 62 kilodalton WT fiber protein.

These results confirm that the method identified herein can be employed to introduce modifications into the fiber protein to produce a chimeric fiber protein. Similar techniques can be employed to introduce modifications into the hexon or penton proteins, or to introduce similar modifications (e.g., the addition of a string of amino acids comprised of arginine, lysine and/or histidine, or comprised of aspartate and/or glutamate, or the addition of any of these sequences into a coding region of the coat proteins).

Example 3

This example describes the binding to cells of an adenoviral vector comprising a chimeric coat protein such as a chimeric fiber protein as compared with a wild-type adenoviral vector, either in the presence or absence of added soluble wild-type fiber protein

For these experiments, the cells identified in Example 1 to which adenovirus binds with either high efficiency (i.e. receptor-plus cells) or low efficiency (i.e., receptor-minus cells) were employed. The epithelial cell line A549 was used as representative of receptor-plus cells, and the fibroblast cell line HS 68 was used as representative of receptor-minus cells. Confluent monolayers of either A549 or HS 68 cells were preincubated at 4°C with concentrations of soluble fiber protein ranging from 0 to about 10 µg/ml. The GV10 UTV vector comprising chimeric fiber protein (UTV) or GV10 vector comprising wild-type fiber protein (WT) were labeled with tritiated thymidine as described in Example

1. About 20,000 cpm of [³H]-thymidine labeled GV10 UTV or GV10 vector were then incubated with the cells for about 2 hours at 4°C. The cells were washed three times with cold PBS, and the cell-associated cpm were determined by scintillation counting. Results were obtained as the average of duplicate measurements and are presented in Figures 6A and 6B for the A549 and HS 68 cell lines, respectively.

As can be seen in Figures 6A-B, the GV10 UTV vector chimeric fiber protein was able to bind both receptor-plus (Fig. 6A) and receptor-minus (Fig. 6B) cells with high efficiency. In comparison, the GV10 vector comprising wild-type fiber was more effective at binding to receptor-plus cells. In particular, radiolabeled GV10 UTV bound to cells expressing detectable levels of fiber receptor (i.e., A549 alveolar epithelial cells) about 2- to 2.5-fold better than GV10. Whereas all of the binding of the GV10 vector was inhibited by competing recombinant fiber protein, only about 40% of the GV10 UTV vector was inhibited by the addition of competing fiber. No detectable binding of GV10 vector comprising wild-type adenoviral fiber to HS 68 human foreskin fibroblast cells lacking fiber receptor was observed. In comparison, the GV10 UTV vector efficiently bound to HS 68 cells, and the addition of competing fiber protein had no effect on binding.

These results confirm that binding of the GV10 UTV vector comprising a chimeric coat protein (i.e., a chimeric fiber protein) does not occur via the wild-type adenoviral fiber receptor, and instead occurs via a heretofore unrecognized fiber receptor. Moreover, the results confirm that incorporation of a chimeric coat protein such as a chimeric fiber protein into an adenoviral vector results in an improved adenoviral vector. Namely, the modification comprised by the GV10 UTV vector enables it overcome the aforementioned relative inability of wild-type adenovirus to bind to

receptor-minus cells, in particular, non-epithelial cells, and also allows the modified vector to bind to receptor-plus cells with an increased efficiency.

5

Example 4

This example describes an investigation of the ability of various soluble factors, and inhibitors of these soluble factors, to block binding of adenovirus comprising chimeric fiber protein to receptor-minus HS 68
10 fibroblast cells.

For these experiments, the inhibition of GV10 UTV binding by various negatively charged molecules including salmon sperm DNA, mucin, chondroitin sulfate, and heparin, was assessed. Chondroitin sulfate and heparin
15 are negatively charged molecules which get their charge from sulfate groups. Mucin is negatively charged due to the presence of sialic acid moieties, and DNA is negatively charged due to its incorporation of phosphate moieties. About 20,000 cpm of UTV in 250 μ l of binding
20 buffer (i.e., Dulbecco's Modified Eagle Media (D-MEM) was incubated at room temperature for about 30 minutes with concentrations of negatively charged molecules ranging from about 1×10^{-3} to about 1×10^4 μ g/ml. Following incubation, the mixtures were chilled on ice, and were
25 then added to prechilled HS 68 cells plated in 24 well plates. The cells were incubated for about 1 hour, and then the cells were washed three times with PBS. Cell-associated cpm were determined by scintillation counting, and reported as the average of duplicate measurements.

30

As indicated in Figure 7, whereas the presence of competing wild-type fiber protein had no effect on binding of a GV10 UTV vector (i.e., comprising chimeric fiber) to HS 68 cells, negatively-charged competing
35 molecules were able to block GV10 UTV binding. All four molecules were able to inhibit GV10 UTV binding to HS 68 cells, although heparin and DNA were most effective. These molecules have no significant effect on the binding

of a GV10 vector (i.e., comprising wild-type fiber) to cells expressing high levels of fiber receptor (i.e., A549 cells; data not shown).

These results confirm that negatively charged molecules are able to block binding of the GV10 UTV vector to cells mediated by chimeric fiber protein. This inhibition presumably is due to the binding of the negatively charged molecules to the positively charged polylysine residues present on the GV10 UTV fiber. Accordingly, the impact of enzymes which cleave these negatively charged molecules on binding to cells of the GV10 UTV vector was assessed.

HS 68 cells were plated in 24 well plates, and were preincubated with the dilutions of heparinase (Sigma, St. Louis, MO), chondroitinase (Sigma), and sialidase (Boehringer Mannheim, Inc.) ranging from about 0.0001 to 1 for 45 minutes at 37°C, followed by 15 minutes at 4°C. Whereas chondroitinase cleaves chondroitin sulfate, heparinase cleaves heparin and heparin sulfate, and sialidase cleaves sialic acid. The initial starting concentrations for dilutions were as follows: heparinase, 25 U/ml (U = 0.1 μ mole/hour, pH = 7.5, 25°C); chondroitinase, 2.5 U/ml (U = 1.0 μ mole/minute, pH = 8.0, 37°C); and sialidase 0.25 U/ml (U = 1.0 μ mole/minute, pH = 5.5, 37°C). Following incubation, the cells were washed three times with cold PBS, and were then incubated with 20,000 cpm of labeled GV10 UTV vector for about 1 hour at 4°C. The cells were then washed three times with cold PBS, and the cell-associated cpm were determined by scintillation counting. The results were reported as the average of duplicate measurements.

As illustrated in Figure 8, pretreatment of HS 68 cells with enzymes that remove negatively charged molecules from the cell surface confirms that the GV10 UTV vector comprising the chimeric fiber protein interacts with negatively charged sites on the cell surface. In particular, heparinase and sialidase were

both able to reduce GV10 UTV binding, although heparinase was more effective than sialidase on HS 68 cells.

Thus, these results confirm that a vector comprising chimeric fiber protein (e.g. a GV10 UTV vector), unlike wild-type adenovirus, interacts in a novel fashion with negatively charged molecules on the cell surface to effect cell entry. These results further demonstrate that a vector comprising negatively charged residues (e.g., aspartate and glutamate) instead of positively charged molecules (e.g., lysine) similarly can be employed to bind to and effect cell entry via positively charged molecules present on the cell surface.

Example 5

This example evaluate gene delivery to different types of cells mediated by an adenoviral vector comprising chimeric coat protein such as chimeric fiber protein (e.g., GV10 UTV) as compared to gene delivery mediated by adenovirus comprising wild-type coat protein such as fiber protein (e.g., GV10).

For these experiments, the relative levels of lacZ gene delivery by a vector containing the wild-type fiber protein (i.e., GV10) as compared with vector containing chimeric fiber protein (i.e., GV10 UTV) were compared in epithelial-like cells (i.e., HeLa, A549, HepG2 and H700 T cells), smooth muscle cells (i.e., HA SMC and HI SMC cells), endothelial cells (i.e., HUVEC and CPAE cells), fibroblast cells (i.e., HS 68 and MRC-5 cells), glioblastoma cells (i.e., U118 cells) and monocyte macrophages (i.e., THP-1 cells). Approximately 2×10^5 cells were inoculated one day prior to transduction by adenovirus into 24 multiwell plates. Each well was then infected at an MOI of 1 with GV10 (i.e., comprising wild-type adenoviral fiber protein) or with GV10 UTV (i.e., comprising chimeric adenoviral fiber) in a 250 μ l volume for about one hour. The wells were then washed and incubated for two days, after which the lacZ activity of

the cell lysates was determined. The results were reported as the average of duplicate measurements.

As illustrated in Figure 9, the use of the GV10 UTV vector to transfer a reporter gene to a panel of cell lines confirms that the presence of the chimeric fiber protein (UTV) increases *lacZ* gene delivery to cells expressing low or undetectable levels of fiber receptor (i.e., receptor-minus cells) from about 5- to about 300-fold as compared with wild-type vector (GV10). In cells expressing high levels of the fiber receptor (i.e., receptor-plus cells), the incorporation of the chimeric fiber protein in the GV10 UTV vector results in an increase in gene delivery of up to about 3-fold.

This reduction in expression observed with transduction of receptor-minus non-epithelial cells as compared with receptor-plus epithelial cells by adenovirus comprising a wild-type fiber protein (i.e., GV10) directly correlates with the relative ability of the vector to bind these different cell types, as reported in Example 3. These results support the view that the low expression of receptors for wild-type adenovirus fiber protein is a significant limiting factor to their efficient transduction by current adenovirus vectors.

Similarly, the ability of the chimeric coat protein (i.e., the chimeric fiber protein) to augment gene transfer *in vivo* was assessed. Three BALB/c mice were inoculated intranasally with about 1×10^8 pfu of GV10 in 50 μ l of a saline solution comprising 10 mM MgCl₂ and 10 mM Tris (pH 7.8). Another three mice received the same dose of GV10 UTV, and two mice received the saline solution alone. The animals were sacrificed at two days post-administration, and the lungs were assayed for *lacZ* activity. The lungs were prepared for analysis by snap-freezing the lung in liquid nitrogen, grinding the tissue with a mortar and pestle, and lysing the ground tissue in 1.0 ml of *lacZ* reporter lysis buffer (Promega Corp.,

Madison, WI). A fluorometric assay was used to monitor lacZ activity, and the results of the experiments were reported as the average activity measured from each group of animals.

5 The results of these experiments are illustrated in Figure 10. As can be seen from this Figure, gene transfer *in vivo* mediated by the GV10 UTV vector comprising chimeric fiber protein (UTV) as compared with a vector comprising wild-type fiber protein (GV10)
10 resulted in an average of 8-fold higher delivery to mouse lung.

These results thus confirm that incorporation of a chimeric coat protein (in this case, a chimeric fiber protein) in an adenoviral vector substantially increases
15 the efficiency of vector-mediated gene delivery both *in vitro* and *in vivo* as compared to an adenovirus vector comprising wild-type fiber protein. Moreover, the results support the conclusion that low fiber receptor expression is a significant factor contributing to the
20 suboptimal delivery observed in the lung and in other tissues. Also, the results confirm the superiority of the GV10 UTV vector, as well as other similar UTV vectors, over other currently available adenoviral vectors for gene transfer (e.g., delivery of the CFTR
25 gene) to the lung and other tissues.

Example 6

This example evaluates the ability of a vector according to the invention comprising a chimeric coat
30 protein (e.g., a chimeric fiber protein) to interact with passenger DNA by means of a protein/DNA interaction, and to thereby carry the DNA into the cell in a "piggy-back" fashion.

For these experiments, an adenoviral vector
35 comprising wild-type fiber (i.e., GV10) and an adenoviral vector comprising chimeric fiber (i.e., GV10 UTV) were used to assess gene transfer to receptor-plus epithelial

cells (i.e., 293, A549, and H700 T cells). In control experiments, the cells were transduced with the vectors as previously described. In the experimental condition, the vectors were incubated with the plasmid pGUS, which
5 comprises a β -glucuronidase reporter gene, such that the chimeric adenoviral fiber protein was able to complex with the plasmid DNA. Specifically, about 5×10^7 active particles (i.e., fluorescence focus units (ffu)) of GV10 or GV10 UTV were incubated for 1 hour with about 2.5 μ g
10 of plasmid pGUS DNA. The mixture was then added to about 2×10^5 of the indicated cells in 250 μ l of DMEM containing 10% fetal bovine serum. Both β -glucuronidase and β -galactosidase activity were then assessed by fluorometric assay at 10 days post-transduction. β -
15 glucuronidase expression in cells was monitored similarly to the β -galactosidase assay for *lacZ* expression, by monitoring the generation of a blue color when β -glucuronidase catalyzes a reaction with the substrate X-glu.

20 The results of these experiments are illustrated in Figure 11. Comparable levels of *lacZ* expression were obtained when either a GV10 vector (i.e. comprising wild-type fiber protein) or a GV10 UTV vector (i.e. comprising chimeric fiber protein) were employed to transfer the
25 reporter gene in *cis* to epithelial cells. In comparison, the wild-type vector was able to transfer intracellularly the plasmid pGUS at only a relatively low level in all epithelial cells, as assessed by β -glucuronidase gene expression. This basal level of gene transfer likely was
30 accomplished by means of receptor-mediated uptake (RME) of bystander molecules, as previously described (PCT patent application WO 95/21259). However, with use of a GV10 UTV vector comprising a chimeric fiber protein, transfer of the pGUS plasmid was substantially increased.
35 In the case of gene transfer to 293 cells, pGUS plasmid-directed β -glucuronidase expression exceeded expression

observed following GV10 UTV-vector mediated transfer of a cis-linked reporter gene.

These results confirm that a vector comprising a chimeric coat protein such as a chimeric fiber protein according to the invention demonstrates increased transfer of a nucleic acid that is not located in *cis* with the vector. Ostensibly, this enhanced gene transfer is effected by the occurrence of a protein/DNA interaction between the negatively charged residues on the chimeric fiber (e.g., residues of the polylysine string), resulting in binding to the vector of the nucleic acid; however, other means of enhancement also are possible.

15 Example 7

This example describes the construction of further plasmids containing UTV or UTV-like sequences in the C-terminus of the fiber protein.

The transfer plasmid, p193(F5*) (Figure 12; also known as p193NS (F5*), pNS (F5*), and pAd NS 83-100 UTV) described in Example 2 was employed as a starting point for the construction of these further plasmids containing chimeric adenovirus fiber proteins. As depicted in Figure 12, p193 (F5*) contains a mutated fiber gene with a BamHI site between the last fiber protein codon and the frameshifted fiber protein stop codon. The further mutant transfer plasmids constructed as described herein contain sequences in the fiber C-terminus encoding an amino acid glycine/serine repeat linker, a targeting sequence, and a stop codon. These plasmids were made by cloning synthetic oligonucleotides into the BamHI site of p193(F5*) to create the transfer plasmid p193NS (F5*) pGS(K7) (also known as p193 (F5*) pGS(K7) or pNS (F5*) pK7) depicted in Figure 13.

35 Thus, the sequence of the wild-type Ad5 fiber gene
is:

TCA TAC ATT GCC CAA GAA TAA AAA AGAA
Ser Tyr Ile Ala Gln Glu

[SEQ ID NO:59]
[SEQ ID NO:60]

wherein "TAA" is a termination codon, and the polyadenylation sequence is emboldened. The C-terminus of the mutated fiber gene present in p193(F5*) is:

- 5 TCA TAC ATT GCC CAA GAA GGA TCC AATAAA GAA [SEQ ID NO:19]
Ser Tyr Ile Ala Gln Glu Gly Ser [SEQ ID NO:20]

- wherein the underlined sequence indicates the mutated *Bam*HI site introduced into the fiber protein, and the polyadenylation sequence is emboldened. In comparison, the amino acid sequence of the C-terminus of the fiber gene present in p193NS (F5*) pGS(K7) is:

- 15 G S G S G S G S G S **KKKKKKK** [SEQ ID NO:22]

- wherein the underlined sequence indicates the mutated *Bam*HI site introduced into the fiber protein, and the emboldened sequence indicates the polylysine string added to the C-terminus. This amino acid sequence is encoded by the nucleid acid sequence: GGA TCA GGA TCA GGT TCA GGG AGT GGC TCT AAA AAG AAG AAA AAG AAG AAG TAA [SEQ ID NO:21], wherein "TAA" is a termination codon.

- The overlapping synthetic oligonucleotides used to make the transfer plasmid p193NS (F5*) pGS(K7) were: pK7s (sense), GA TCA GGA TCA GGT TCA GGG AGT GGC TCT AAA AAG AAG AAA AAG AAG AAA TAA G [SEQ ID NO:61]; pK7a (antisense), GA TCC TTA CTT CTT CTT TTT CTT CTT TTT AGA GCC ACT CCC TGA ACC TGA TCC T [SEQ ID NO:62]. The sense and antisense oligonucleotides were mixed in equimolar ratios and cloned into the *Bam*HI site of p193NS (F5*) to create p193NS (F5*) pGS(pK7). Verification of the correctly-oriented insert in p193NS (F5*) pGS(pK7) was performed by PCR using the pK7s sense primer and the downstream antisense oligonucleotide primer A5a32938, CAGGTTGAATACTAGGGTTCT [SEQ ID NO:63]. The plasmid was also verified to contain the correctly oriented insert by sequencing the DNA sequence in the region of the insert using the A5a32938 primer.

The transfer plasmid p193NS (F5*) was employed in the construction of further mutant transfer plasmids that additionally contain a UTV or UTV-like cell targeting sequence in the C-terminus of the fiber protein. These
 5 plasmids include p193NS (F5*) pGS(null) (also known as p193 (F5*) pGS(null) or p193 (F5*) pGS), pBSS 75-100 pGS(null), pBSS 75-100 pGS(RK32), pBSS 75-100 pGS(RK33), and pBSS 75-100 pGS(tat).

To construct p193NS (F5*) pGS(null), the
 10 complementary overlapping oligonucleotides pGSs, GATCCGGTTCAGGATCTGGCAGTGGCTCGACTAGTTAAA [SEQ ID NO:64], and pGSa, GATCTTTAACTAGTCGAGCCACTGCCAGATCCTGAACCG [SEQ ID NO:65] were constructed for direct ligation into the *Bam*HI-digested p193NS (F5*) plasmid. Verification of the
 15 correctly-oriented clone was performed by PCR using the pGSs primer and the downstream antisense oligonucleotide primer A5a32938. The plasmid was also verified to contain the correctly oriented insert by sequencing the
 20 DNA sequence in the region of the insert using the A5a32938 primer.

The vector pBSS 75-100 pGS(null) (also known as pBSS 75-100 Δ E3 pGS(null)) depicted in Figure 14 was constructed by replacing the *Nhe*I to *Sal*I fragment from pBSS 75-100 with the corresponding fragment from p193NS
 25 (F5*) pGS(null). The *Spe*I site that is not within the fiber chimera gene was then eliminated by partially restricting the plasmid with *Spe*I, filling in with Klenow fragment and then religating the vector. The resultant vector comprises the relevant nucleic acid sequence:
 30 GCCCAAGAAGGATCCGGTTCAGGATCTGGCAGTGGCTCGACTAGTTAA [SEQ ID NO:23] (wherein "TAA" is a termination codon), which codes for the amino acid sequence Ala Gln Glu Gly Ser Gly Ser Gly Ser Gly Ser Thr Ser [SEQ ID NO:24].

35 The inserts of plasmids pBSS 75-100 pGS(RK32) (also known as pBSS 75-100 Δ E3 pGS(RKKK), or pBSS 75-100 Δ E3 pGS(RKKK2)), pBSS 75-100 pGS(RK33) (also known as pBSS

75-100 Δ E3 pGS(RKKK), or pBSS 75-100 Δ E3 pGS(RKKK3)), and pBSS 75-100 pGS(tat), were constructed for direct ligation into the *Spe*I-digested pBSS 75-100 pGS(null) plasmid. To construct pBSS 75-100 pGS(RK32) depicted in

5 **Figure 15**, the complementary overlapping oligonucleotides, RK32s, CTAGAAAGAAGAAACGCAAAAAGAAGA [SEQ ID NO:66] and RK32a, CTAGTCTTCTTTTTCGTTTCTTCTTT [SEQ ID NO:67] were employed. The resultant vector comprises the relevant nucleic acid sequence:

10 GCCCAAGAAGGATCCGGTTCAGGATCTGGCAGTGGCTCGACTAGAAAGAAGAAACGC AAAAAGAAGACTAGTTAA [SEQ ID NO:25] (wherein "TAA" is a termination codon), which codes for the amino acid sequence Ala Gln Glu Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Thr Arg Lys Lys Lys Arg Lys Lys Lys Thr Ser

15 [SEQ ID NO:26].

To construct pBSS 75-100 pGS(RK33) depicted in **Figure 16**, the complementary overlapping oligonucleotides, RK33s, CTAGAAAGAAGAAGCGCAAAAAAAAAAAGAAGAAGAAGA [SEQ ID NO:68]

20 and RK33a, CTAGTCTTCTTCTTTCTTTTTTTTTTTCGCTTCTTCTTCTTT [SEQ ID NO:69] were employed. The resultant vector comprises the relevant nucleic acid sequence:

GCCCAAGAAGGATCCGGTTCAGGATCTGGCAGTGGCTCGACTAGAAAGAAGAAGCGC AAAAAAAAAAGAAGAAGAAGACTAGTTAA [SEQ ID NO:27] (wherein

25 "TAA" is a termination codon), which codes for the amino acid sequence Ala Gln Glu Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Thr Arg Lys Lys Lys Arg Lys Lys Lys Arg Lys Lys Lys Thr Ser [SEQ ID NO:28].

To construct pBSS 75-100 pGS(tat) the complementary

30 overlapping oligonucleotides, TATs, CT AGT TAT GGG AGA AAA AAG CGC AGG CAA CGA AGA CGG GCA T [SEQ ID NO:70] and TATa, CT AGA TGC CCG TCT TCG TTG CCT GCG CTT TTT TCT CCC ATA A [SEQ ID NO:71] were employed. The resultant vector comprises the relevant nucleic acid sequence: ACT AGT TAT

35 GGG AGA AAA AAG CGC AGG CAA CGA AGA CGG GCA TCT AGT [SEQ ID NO:72], which codes for the amino acid sequence Thr

Ser Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Ala Ser
Ser [SEQ ID NO:73].

Verification of the correctly-oriented clone was performed by PCR using the sense primers (RK32s, RK33s, or TATs) for each of the three respective plasmids, and using the downstream antisense oligonucleotide primer A5a32938. Each of the plasmids also were verified to contain the correctly-oriented insert by sequencing the DNA sequence in the region of the insert using the A5a32938 primer.

Example 8

This example described the construction of plasmids containing UTV domains in the fiber loop.

Plasmids containing a UTV sequence (and/or a spacer sequence) in an exposed loop of the fiber protein are constructed by incorporating any of the aforesaid sequences (as well as any further UTV-like sequences) into the fiber protein. This is accomplished making use of the plasmid transfer vector p193NS (F5*) to construct the further transfer vector p193NS F5F2K (also called p193 F5F2K) depicted in Figure 17. Plasmid p193NS F5F2K contains a unique *Spe* I restriction site within the Ad2 fiber gene encoding an exposed loop in the protein. Namely, the fiber gene present in p193NS F5F2K comprises the fiber sequence:

ATT ACA CTT AAT GGC ACT AGT GAA TCC ACA
Ile Thr Leu Asn Gly Thr Ser Glu Ser Thr

GAA ACT [SEQ ID NO:29]
Glu Thr [SEQ ID NO:30]

wherein the underlined sequence indicates the novel *Spe* I site introduced into the fiber gene.

This vector was then used to clone targeting sequences into the *Spe* I site. In particular, a nucleic acid sequence encoding the stretch of 8 basic amino acids RKKKRKKK (Arg Lys Lys Lys Arg Lys Lys Lys [SEQ ID NO:74]) comprising the heparin binding domain were cloned into

the *Spe* I site of p193 F5F2K using overlapping sense and antisense oligonucleotides.

Namely, the (RKKK)₂ sequence comprises, in part, the sequence:

5 TCT AGA AAA AAA AAA CGC AAG AAG AAG ACT
Thr Arg Lys Lys Lys Arg Lys Lys Lys Thr

AGT
Ser

[SEQ ID NO:75]

[SEQ ID NO:76].

10 The 27-mer sense oligonucleotide RK32s and 27-mer
antisense oligonucleotide RK32a described in Example 7
were employed for cloning the PolyGS(RKKK)₂ sequence
comprising the RKKKRKKK [SEQ ID NO:74] peptide motif.
The p193NS F5F2K(RKKK)₂ plasmid was constructed by cloning
15 the DNA sequence encoding the binding domain into the *Spe*
I site of p193NS F5FK2. The overlapping sense and
antisense oligonucleotides encoding the binding domain
were first annealed and then directly ligated into the
Spe I restriction site to result in the plasmid p193NS
20 F5F2K(RKKK)₂, depicted in Figure 18. This plasmid also is
known as p193NS F5F2K(RKKK2), p193NS F5F2K(RK32), or p193
F5F2K(RKKK2). The relevant portion of the modified loop
of the fiber knob present in p193NS F5F2K(RKKK)₂ is:

25 ATT ACA CTT AAT GGC ACT AGA AAG AAG AAA CGC AAA AAG AAG
Ile Thr Leu Asn Gly Thr Arg Lys Lys Lys Arg Lys Lys Lys

ACT AGT GAA TCC ACA GAA ACT
Thr Ser Glu Ser Thr Glu Thr

[SEQ ID NO:31]

[SEQ ID NO:32].

30 Furthermore, a (RKKK)₂ sequence, or other variations
of this sequence, can be inserted into p193NS F5F2K.

This sequence comprises, in part:

TCT AGA AAG AAG AAG CGC AAA AAA AAA AGA AAG AAG AAG
Thr Arg Lys Lys Lys Arg Lys Lys Lys Arg Lys Lys Lys

35 ACT AGT
Thr Ser

[SEQ ID NO:77]

[SEQ ID NO:78].

The sequence can be inserted with use of the 39-mer sense
oligonucleotide (RKKK)₂(s) (i.e., comprising the sequence
CT AGA AAG AAG AAG CGC AAA AAA AAA AGA AAG AAG AAG A [SEQ
40 ID NO:79]), and the 39-mer antisense oligonucleotide
(RKKK)₂(a) (i.e., comprising the sequence CT AGT CTT CTT
CTT TCT TTT TTT TTT GCG CTT CTT CTT T [SEQ ID NO:80]).

The resultant plasmid p193NS F5F2K(RKKK), is depicted in Figure 19. This plasmid also is known as p193NS F5F2K(RKKK3), p193 F5F2K(RKKK3), or p193 F5FK(RK33). The relevant portion of the modified loop of the fiber knob present in p193 F5F2K(RKKK), is:

CTT AAT GGC ACT AGA AAG AAG AAG CGC AAA AAA AAA AGA AAG
Leu Asn Gly Thr Arg Lys Lys Lys Arg Lys Lys Lys Arg Lys

AAG ACT AGT GAA TCC ACA [SEQ ID NO:33]
10 Lys Thr Ser Glu Ser Thr [SEQ ID NO:34].

Example 9

This example describes the construction of plasmids containing chimeric penton base proteins comprising UTV or UTV-like sequences.

The transfer plasmid pACT (Δ RGD) (also described as plasmid pAT in U.S. Patent 5,559,099) was derived, in part, by manipulating a plasmid containing the unique BamHI/PmeI fragment (13259-21561) of the Ad5 genome, and contains, among other things, a penton base protein comprising a deletion of 8 amino acids constituting the α , integrin binding domain, and a substitution of the deleted region for amino acids constituting a unique SpeI site, for the convenient insertion of exogenous sequences.

To construct plasmid pACT (RKKK), (also known as pACT (RKKK3) or pACT (RK33)) depicted in Figure 20, the complementary overlapping oligonucleotides RK33s and RK33a were directly ligated into a SpeI-digested pACT (Δ RGD) plasmid. Verification of the correctly-oriented clone was performed by PCR using the RK33a primer for the plasmid and the upstream sense oligonucleotide primer A5s15002. The plasmid also was verified to contain the correctly oriented insert by sequencing the DNA sequence in the region of the insert using the A5s15002 primer. The relevant portion of the UTV domain that will result in the chimeric penton base protein in pACT (RKKK), is:

AAC GAT ACT AGA AAG AAG AAG CGC AAA AAA AAA AGA AAG AAG AAG

Asn Asp Thr Arg Lys Lys Lys Arg Lys Lys Lys Arg Lys Lys Lys

ACT AGT GCC ACA [SEQ ID NO:35]

Thr Ser Ala Thr [SEQ ID NO:36] .

5 The plasmid pACT (RK32) (which also can be called
pACT (RKKK2) or pACT (RKKK)₂) depicted in **Figure 21**
similarly can be constructed using the RK32s and RK32a
overlapping primers. The relevant portion of the UTV
domain present in the chimeric penton base protein in
10 pACT (RKKK)₂ is:

AAC GAT ACT AGA AAG AAG AAG AGA AAG AAG AAG ACT

Asn Asp Thr Arg Lys Lys Lys Arg Lys Lys Lys Thr

AGT GCC ACA [SEQ ID NO:37]

15 Ser Ala Thr [SEQ ID NO:38] .

Example 10

This example describes the construction of plasmids
containing UTV or UTV-like sequences in the adenovirus
20 hexon protein, and particularly which contain these
sequences in an exposed loop of the adenovirus hexon
protein.

These plasmids can be constructed making use of
another transfer plasmid, plasmid pACT H11, depicted in
25 **Figure 22**. Plasmid pACT H11 itself is derived from
plasmid pACT (comprising from 13259-21561 of the Ad5
genome), which contains the majority of the hexon protein
coding sequence (corresponding to about 18842-21700). In
particular, pACT H11 can be constructed by incorporating
30 an *Xba*I site into the loop 1 region of the Ad5 hexon
protein. Similar techniques can be used to incorporate
an *Xba*I site, or any other convenient restriction site,
into either the loop 1 or the loop 2 region, or into
another exposed loop of the hexon protein. Sense and
35 antisense primers can be used to amplify the loop 1
region from Ad5 DNA by PCR, and at the same time
introduce a mutation which results in a unique mutated
*Xba*I site in the loop 1 region.

In particular, the sense primer,

40 GGACAGGGGCCCTACTTTTAAGCCCTACTCTGGCA [SEQ ID NO:81],

containing the naturally occurring unique restriction site *ApaI* that occurs in pACT, and the antisense primer, ATCTTCACTGTACAATACCACTTTAGGAGTCAAGTTATCACCTCTAGATGCGGTCGC CT [SEQ ID NO:82], containing the unique restriction site, *BsrGI*, can be employed. The PCR product, containing the *XbaI* site can then be cut with *BsrGI* and *ApaI*, and cloned back in to pACT to replace the *ApaI* to *BsrGI* fragment. The resultant plasmid, pACT H11, contains a unique *XbaI* site for the insertion of UTV sequences into loop 1 of the hexon. The presence of the *XbaI* site in the pACT H11 clone can be verified by restriction digestion using *XbaI*, which should linearize the plasmid.

Part of the unmutated hexon loop 1 amino acid sequence comprises the sequence TEATGNGDNL [SEQ ID NO:83]. In comparison, the mutated hexon loop 1 amino acid sequence following the wild-type TEA residues in pACT H11 (Figure 22) comprises the sequence TASRGDNL [SEQ ID NO:40] (i.e., encoded by the nucleic acid sequence ACCGCATCTAGAGGTGATAACTTG [SEQ ID NO:39]).

The *XbaI* site of pACT H11 then can be used as a unique site in which to clone universal targeting sequences such as RKKKRKKK [SEQ ID NO:74], for instance, using the overlapping oligonucleotides, RK32s and RK32a. The particular plasmid that results from such manipulations, i.e. pACT H11 (RKKK), (or pACT H11 (RK32) or pACT H11 (RKKK2)) is depicted in Figure 23. This plasmid comprises the sequence:

ACC GCA TCT AGA AAG AAG AAA CGC AAA AAG AAG ACT AGA
Thr Ala Ser Arg Lys Lys Lys Arg Lys Lys Lys Thr Arg
GGT GAT AAC TTG [SEQ ID NO:41]
Gly Asp Asn Leu [SEQ ID NO:42].

Other UTV or UTV-like sequences also can be cloned in the loop 1 region of the hexon protein, and/or into the loop 2 region of the hexon protein. For instance, a similar approach can be used to mutate the sequence encoding the hexon loop 2 to make plasmid pACT H12 (not shown) that contains a unique restriction site (such as a

XbaI site) into which further UTV sequences can be cloned.

Also, plasmid pACT H11 (RKKK), (or pACT H11 (RKKK3) or pACT H11 (RK33)) can be constructed making use of the complementary overlapping oligonucleotides RK33s and RK33a, and directly ligating the PCR product into the XbaI-digested pACT H11 plasmid. Verification of the correctly-oriented clone can be performed by PCR using the RK32 sense primer for the plasmid and an appropriate downstream antisense oligonucleotide primer. The plasmid can be verified to contain the correctly oriented insert by sequencing the DNA sequence in the region of the insert using the downstream antisense primer. Similar approaches can be employed for construction of analogous transfer vectors, particularly the analogous transfer vectors pACT H12 (RKKK), and pACT H12 (RKKK),.

Example 11

This example describes the construction of a plasmid having a short-shafted fiber protein. In particular, this example describes the construction of the plasmid, p193 F5F9sK.

The plasmid p193 F5F9sK (also known as p193 F5F9K-Short) is depicted in Figure 24. This vector encodes a chimeric fiber protein wherein approximately two thirds of the Ad5 fiber shaft is deleted and the Ad5 fiber knob is replaced with the Ad9 fiber knob.

The plasmid p193F5F9K-short was constructed from p193NS (F5*). The oligonucleotide primers GGACTAGTAG CATTTAATAA AAAAGAAGAT AAGCGC [SEQ ID NO:84] and CCGGATCCTC ATTCTTGGGC GATATAGG [SEQ ID NO:85] were used to amplify the Ad9 sequence encoding the last shaft repeat and knob from the fiber gene. The PCR product was then purified, using standard techniques, and digested with the restriction enzymes *NheI* and *BamHI*, which allowed cloning of the PCR product into the *NheI/BamHI* region of the p193NS (F5*) transfer plasmid. The

resultant short-shafted fiber protein can be employed in construction of adenoviral vectors, as described below. Furthermore, any one or more of the aforementioned UTV or UTV-like sequences can be incorporated into the short-shafted fiber, and the resultant fiber can be employed for cell delivery.

Example 12

This example describes the construction of plasmids containing UTV or UTV-like sequences in an extended structure, particularly in hexon and/or penton base protein, so as to result in lengthened hexon and/or penton base proteins that accordingly are better able to contact cells and participate in cell targeting. The resultant chimeric proteins are "spiked" in the sense that they comprise an insertion of a nonnative amino acid sequence that will jut out from the virus surface.

The primers 1alpha(s),

GGGCTGCAGGCGGCCGAGAGCTGAAGAGGCAGCCACACGGGCTGAGGAGAA [SEQ ID NO:86], and 1alpha(a),

GGGGTGCACACAGCTTCGGCCTTAGCGTTAGCCTGTTTCTTCTGAGGCTTCTCGACCT [SEQ ID NO:87], can be used to amplify the region of the penton gene encoding the 32 amino acid α -helical domain that follows the RGD sequence. This 32 amino acid sequence comprises the sequence

ATRAEEDRAEAEAAAEAAAPAAQPEVEKPQKK [SEQ ID NO:88]. The primers used also can encode an additional α -helical sequence on either end, such that, for instance, the final amplified DNA sequence encodes the sequence:

CTG CAG GCG GCC GCA GAA GCT GAA GAG GCA GCC ACA CGG GCT GAG
Leu Gln Ala Ala Ala Glu Ala Glu Glu Ala Ala Thr Arg Ala Glu
 GAG AAG CGC GCT GAG GCC GAA GCA GCG GCC GAA GCT GCC GCC CCC
 Glu Lys Arg Ala Glu Ala Glu Ala Ala Ala Glu Ala Ala Ala Pro
 GCT GCG CAA CCC GAC GTC GAG AAG CCT CAG AAG AAA CAG GCT AAC
 Ala Ala Gln Pro Glu Val Glu Lys Pro Gln Lys Lys Gln Ala Asn
 GCT AAG GCC GAA GCT GTG CAG GCG GCC GCA GAA GCT GAA GAG GCA
 Ala Lys Ala Glu Ala Val Gln Ala Ala Ala Glu Ala Glu Glu Ala

GCC ACA CGG GCT GAG GAG AAG CGC GCT GAG GCC GAA GCA GCG GCC
Ala Thr Arg Ala Glu Glu Lys Arg Ala Glu Ala Glu Ala Ala Ala

5 GAA GCT GCC GCC CCC GCT GCG CAA CCC GAG GTC GAG AAG CCT CAG
Glu Ala Ala Ala Pro Ala Ala Gln Pro Glu Val Glu Lys Pro Gln

AAG AAA CAG GCT AAC GCT AAG GCC GAA GCT GTG CAC
[SEQ ID NO:89]

10 Lys Lys Gln Ala **Asn** Ala Lys Ala Glu Ala Val His
[SEQ ID NO:90]

where the emboldened sequence corresponds to the
nonpenton sequence encoded by the primers, and the
underlined sequence represents the amino acids encoded by
15 the two compatible restriction sites, *SfcI* and *ApaLI*.
These amino acids also preserve the integrity of an alpha
helix according to standard computer programs designed to
predict α helix structure.

The PCR product encoding these amino acids can be
20 cut with both *SfcI* and *ApaLI*, religated, and then cut
again with both enzymes. Ligation of like-site to like-
site preserves the site for recutting; however, ligation
of the compatible, but unlike sites, destroys the site.
Therefore, upon recutting of the ligated product,
25 multiple fragments will be produced which are multiples
of the original size of the PCR product. There will be
completely recut fragments (approximately 150 bp),
approximately 300 bp fragments (having one restriction
site destroyed), and approximately 450 bp fragment
30 (having 2 sites destroyed), and so on. The procedure
accordingly allows the original sequence encoding 50
amino acids (1alpha) to be doubled (2alpha), tripled
(3alpha), and so on, for cloning large, uninterrupted α -
helical regions into a protein to create a larger
35 "spike", or extension of the protein.

For instance, the 2alpha double product (i.e.,
2alpha2) can be cloned into the first *Ppu10I* site of the
plasmid, pSPdelta (depicted in **Figure 25**), to create the
plasmid pSP2alpha (depicted in **Figure 26**). The plasmid
40 pSPdelta is constructed from the base plasmid pUC19, or
any other suitable cloning plasmid. The pSPdelta

transfer plasmid can be employed to make further modifications of the penton or hexon protein that allow the UTV sequence (or any other targeting sequence) to be elevated out from the virion surface. This elevation of the targeting sequence in a spiked structure (e.g., a type of "tower") will minimize steric hindrance of the penton and hexon interaction with the cell surface by the fiber protein, and will allow greater access of the chimeric penton and hexon proteins for interacting with the cell surface.

The plasmid pSPdelta contains a unique *SpeI* cloning site for the incorporation of UTV sequences. The *SpeI* cloning site is flanked on either side by *Ppu10I* cloning sites which allow the incorporation of DNA sequences encoding amino acids that will elevate the UTV sequence away from the virion surface. The pSPdelta plasmid was constructed by cloning into the unique *XbaI* site of pUC19 overlapping oligonucleotides which are designed to insert directly into the *XbaI* site. In particular, the sense oligonucleotide is:

CTAGAGCAGCTATGCATGAAGGGACTAGTGGAGAGATGCATGCAGCCT [SEQ ID NO:91]. The antisense, complementary oligonucleotide is: CTAGAGGCTGCATGCATCTCTCCACTAGTCCCTTCATGCATAGCTGCT [SEQ ID NO:92]. The oligonucleotides are mixed in equimolar ratios and cloned into the *XbaI* site of pUC19. The presence of the correct insert can be confirmed by sequencing across the region of the insert, and by cutting the plasmid with *Ppu10I*. The *XbaI* sites on either side of the insert allow the convenient removal of this section in later clones.

Because there are two *Ppu10I* sites in pSPdelta, the plasmid can be partially restricted upon digestion with *Ppu10I* so that only a single site is cut. Ligation of the PCR product comprising the *ApaI* and *SfcI* sites into the compatible *Ppu10I* site will destroy the first *Ppu10I* site in the plasmid and also destroy the *ApaI* and *SfcI* sites. These destroyed sites are represented in Figure

26 (showiing the pSP2alpha2 plasmid) by a "j". A second
 doublet product then can be cloned into the remaining
 Ppu10I site of the pSP2alpha plasmid to produce the
 plasmid pSP2alpha2 depicted in Figure 27, which contains
 5 an SpeI site for insertion cloning of a UTV sequence, or
 other similar targeting sequence. Computer analysis of
 the secondary structure of the anticipated 2alpha2
 protein confirms that it will be a complete α helix
 except for in the region of the SpeI cloning site.

10 Thus, whereas pSPdelta comprises the sequence:

TCT AGA GCA GCT ATG CAT GAA GGG ACT AGT GGA GAC ATG CAT GCA
 Ser Arg Ala Ala Met His Glu Gly Thr Ser Gly Glu Met His Ala

15 GCC TCT AGA [SEQ ID NO:43]
 Ala Ser Arg [SEQ ID NO:44],

pSP2alpha comprises the sequence:

20 TCT AGA GCA GCT ATG CAG GCG GCC GCA GAA GCT GAA GAG GCA
 Ser Arg Ala Ala Met Gln Ala Ala Ala Glu Ala Glu Glu Ala

GCC ACA CGG GCT GAG GAG AAG CGC GCT GAG GCC GAA GCA GCG
 Ala Thr Arg Ala Glu Glu Lys Arg Ala Glu Ala Glu Ala Ala

25 GCC GAA GCT GCC GCC CCC GCT GCG CAA CCC GAG GTC GAG AAG
 Ala Glu Ala Ala Ala Pro Ala Ala Gln Pro Glu Val Glu Lys

CCT CAG AAG AAA CAG GCT AAC GCT AAG GCC GAA GCT GTG CAG
 Pro Gln Lys Lys Gln Ala Asn Ala Lys Ala Glu Ala Val Gln

30 GCG GCC GCA GAA GCT GAA GAG GCA GCC ACA CGG CCT GAG GAG
 Ala Ala Ala Glu Ala Glu Glu Ala Ala Thr Arg Ala Glu Glu

35 AAG CGC GCT GAG GCC GAA GCA GCG GCC GAA GCT GCC GCC CCC
 Lys Arg Ala Glu Ala Glu Ala Ala Ala Glu Ala Ala Ala Pro

GCT GCG CAA CCC GAG GTC GAG AAG CCT GAG AAG AAA CAG GCT
 Ala Ala Gln Pro Glu Val Glu Lys Pro Gln Lys Lys Gln Ala

40 AAC GCT AAG GCC GAA GCT GTG CAT GAA GGG ACT AGT GGA GAG
 Asn Ala Lys Ala Glu Ala Val His Glu Gly Thr Ser Gly Glu

45 ATG CAT GCA GCC TCT AGA [SEQ ID NO:45]
 Met His Ala Ala Ser Arg [SEQ ID NO:46],

and pSP2alpha2 comprises the sequence:

TCT AGA GCA GCT ATG CAG GCG GCC GCA GAA GCT GAA GAG GCA GCC
 Ser Arg Ala Ala Met Gln Ala Ala Ala Glu Ala Glu Glu Ala Ala

ACA CGG GCT GAG GAG AAG CGC GCT GAG GCC GAA GCA GCG GCC GAA
 Thr Arg Ala Glu Glu Lys Arg Ala Glu Ala Glu Ala Ala Glu
 5 GCT GCC GCC CCC GCT GCC CAA CCC GAG GTC GAG AAG CCT CAG AAG
 Ala Ala Ala Pro Ala Ala Gln Pro Glu Val Glu Lys Pro Gln Lys
 AAA CAG GCT AAC GCT AAG GCC GAA GCT GTG CAG GCG GCC GCA GAA
 Lys Gln Ala Asn Ala Lys Ala Glu Ala Val Gln Ala Ala Ala Glu
 10 GCT GAA GAG GCA GCC ACA CGG GCT GAG GAG AAG CGC GCT GAG GCC
 Ala Glu Glu Ala Ala Thr Arg Ala Glu Glu Lys Arg Ala Glu Ala
 GAA GCA GCG GCC GAA GCT GCC GCC CCC GCT GCG CAA CCC GAG GTC
 Glu Ala Ala Ala Glu Ala Ala Ala Pro Ala Ala Gln Pro Glu Val
 15 GAG AAG CCT CAG AAG AAA CAG GCT AAC GCT AAG GCC GAA GCT GTG
 Glu Lys Pro Gln Lys Lys Gln Ala Asn Ala Lys Ala Glu Ala Val
 CAT GAA GGG ACT AGT GGA GAG ATG CAG GCG GCC GCA GAA GCT GAA
 20 His Glu Gly Thr Ser Gly Glu Met Gln Ala Ala Ala Glu Ala Glu
 GAG GCA GCC ACA CGG GCT GAG GAG AAG CGC GCT GAG GCC GAA GCA
 Glu Ala Ala Thr Arg Ala Glu Glu Lys Arg Ala Glu Ala Glu Ala
 25 GCG GCC GAA GCT GCC GCC CCC GCT GCG CAA CCC GAG GTC GAG AAG
 Ala Ala Glu Ala Ala Ala Pro Ala Ala Gln Pro Glu Val Glu Lys
 CCT CAG AAG AAA CAG GCT AAC GCT AAG GCC GAA GCT GTG CAG GCG
 Pro Gln Lys Lys Gln Ala Asn Ala Lys Ala Glu Ala Val Gln Ala
 30 GCC GCA GAA GCT GAA GAG GCA GCC ACA CGG GCT GAC GAG AAG CGC
 Ala Ala Glu Ala Glu Glu Ala Ala Thr Arg Ala Glu Glu Lys Arg
 GCT GAG GCC GAA GCA GCG GCC GAA GCT GCC GCC CCC GCT GCG CAA
 35 Ala Glu Ala Glu Ala Ala Ala Glu Ala Ala Ala Pro Ala Ala Gln
 CCC GAG GTC GAG AAG CCT CAG AAG AAA CAG GCT AAC GCT AAG GCC
 Pro Glu Val Glu Lys Pro Gln Lys Lys Gln Ala Asn Ala Lys Ala
 40 GAA GCT GTG CAT GCA GCC TCT AGA [SEQ ID NO:47]
 Glu Ala Val His Ala Ala Ser Arg [SEQ ID NO:48].

Targeting sequences such as UTV or UTV-like
 sequences can be cloned into the *SpeI* site of the plasmid
 45 pSP2alpha2. In particular, the RK32s and RK32a
 overlapping oligonucleotides can be cloned into the *SpeI*
 site to create pSP2alpha2 (RKKK), (or, pSP2alpha2 (RK32)
 or pSPSalpha2 (RKKK2)). The plasmid pSP2alpha2 (RKKK),
 (or, pSP2alpha2 (RK33) or pSPSalpha2 (RKKK3)) can be
 50 similarly constructed. Alternately, the entire 2alpha2
 α -helical domain can be removed from the plasmid by
 restriction with *XbaI* and cloned into the compatible *SpeI*
 site of pACT (Δ RGD) to create pACT 2alpha2 (RKKK), (which

also can be called pACT 2alpha2 (RKKK2) or pACT 2alpha2 (RK32)). Similar techniques can be employed to produce pACT 2alpha2 (RKKK), (which also can be called pACT 2alpha2 (RKKK3) or pACT 2alpha2 (RK33)).

5 Similarly, chimeric hexon proteins that are spiked (i.e., comprise sequences resulting in their extension) can be constructed by cloning the 2alpha2 α -helical domain into the XbaI site of pACT H11 to create pACT H11 2alpha2 (RKKK), (which also can be called pACT H11 2alpha2 (RKKK2) or pACT H11 2alpha2 (RK32)). Similar techniques
10 can be employed to produce pACT H12 2alpha2 (RKKK), (which also can be called pACT H12 2alpha2 (RKKK2) or pACT H12 2alpha2 (RK32)).

15 Example 13

This example describes the construction of further adenoviral vectors, in addition to those previously described, which contain UTV or UTV-like sequences in the adenoviral fiber protein.

20 Construction of adenovirus vectors containing UTV modifications in the fiber can be accomplished in multiple ways by those skilled in the art. One method to create the UTV fiber vectors from plasmids described above is to first linearize the plasmid DNA with SalI and
25 then transfect this DNA into a 293 packaging cell line that was infected just prior to transfection with an E4-deleted adenovirus. E4-deleted adenovirus vectors are incapable of replicating in cell lines such as the 293 cell line, which only provide the adenovirus E1 regions
30 in trans. Recombination of the plasmids which contain the modified fiber gene and the E4 regions with the E4-deleted DNA results in a replication-competent, E4-containing vector which carries the modified fiber gene.

Accordingly, the plasmids p193NS (F5*) pGS(K7), pBSS
35 75-100 pGS(null), pBSS 75-100 pGS(RKKK), pBSS 75-100 pGS(RKKK), pBSS 75-100 pGS(tat), p193NS F5F2K(RK32), and p193 F5F9sK were each linearized with SalI and

transfected into 293 cells infected 1 hour prior with either the E4-deleted adenovirus vector, GV11A.Z (which carries the LacZ gene under the control of a cytomegalovirus (CMV) promoter), or GV11A.S (which carries a secretory alkaline phosphatase gene under the control of a CMV promoter). The resultant adenovirus vectors, AdZ.F(pK7), AdZ.F(pGS), AdZ.F(RKKK), (also known as AdZ.F(RKKK2) or AdZ.F(RK32)), AdZ.F(RKKK), (also known as AdZ.F(RKKK3) or AdZ.F(RK33)), AdZ.F(tat), AdZ.F5F2K(RKKK), (also known as AdZ.F5F2K(RKKK2) or AdZ.F5F2K(RK32)), and AdZ.F5F9sK (also known as AdZ.F5F9K-Short) were obtained and were purified through two successive rounds of plaquing on 293 cells.

All the vectors were verified to contain the correct sequence through PCR across the region of the insert, and by restriction analysis of viral DNA obtained from vector-infected 293 cells by Hirt extraction. Western analysis (as previously described) also can be employed to examine protein size, if so desired. Western analysis of fiber protein from vector particles and/or vector-infected cell lysates electrophoresed on a polyacrylamide gel should show a corresponding shift in the mobility of the fiber protein compared to unmodified fiber protein that is consistent with the presence of additional amino acid sequences. For instance, Western analysis of AdZ.F(pK7) particles verified that its fiber protein is shifted up compared to that of the AdZ vector comprising unmodified fiber, consistent with the presence of additional amino acids in the AdZ.F(pK7) fiber protein.

Other plasmid maps depicted herein similarly can be made into adenoviral vectors by utilizing the same procedure outlined above (or minor variations thereof).

Example 14

This example describes the construction of adenoviral vectors which contain UTV or UTV-like sequences in the adenoviral penton base protein.

The method of making an adenoviral vector comprising a chimeric penton base protein is described, for instance, in Wickham et al., *J. Virol.*, 70, 6831-6838 (1996). A pACT vector described above containing the

5 chimeric penton base protein (e.g., transfer plasmids pACT 2alpha2 (RKKK)₂, pACT H12 2alpha2 (RKKK)₂, and pACT (ΔRGD)) can be digested, for instance, with *Bam*HI, to linearize the plasmid. Ad5 DNA can be digested with the restriction endonuclease *Xmn*I, which cuts wild-type Ad5

10 at positions 14561 and 15710 within the Ad5 genome. The two larger fragments are purified away from the smaller 1 kb piece then transfected along with the linearized plasmid into the appropriate cell line (e.g., a 293 cell line) to produce recombinant virus.

15 Adenoviral vectors produced in this fashion are purified from potentially-contaminating unmodified vectors through two successive rounds of plaque purification on 293 cells. The resultant vectors are then verified to contain the correct sequence in the

20 penton base region through restriction analysis of viral DNA obtained following Hirt extraction of vector-infected 293 cells. Sequencing of PCR products generated by amplifying the region of the insert from the viral DNA can also be used to verify the presence of the insert.

25 Western analysis of the chimeric penton base electrophoresed on a polyacrylamide gel should show a corresponding shift in the mobility of the penton base compared to unmodified penton base that is consistent with the presence of additional amino acid sequences in

30 the chimeric protein.

Example 15

This example describes the construction of adenoviral vectors containing UTV or UTV-like sequences

35 in the hexon protein.

To construct the virus AdZ.H(RKKK)₂, left and right vector arms are prepared that contain DNA sequences which

overlap and will recombine on either side of the *PmeI* to *BamHI* pACT H11(RK32) sequence to create an intact AdZ.H (RK32) genome. To construct the left arm, purified Ad5 DNA is restriction digested with *AgeI*, which cuts Ad5 at positions 14499, 15283, 19017, 23063, 23656, 23411, and 31102. The 0-14499 fragment then can be employed as the left arm and purified from the other fragments by gel electrophoresis. The right arm can be prepared by digesting Ad5 DNA with *DrdI*. *DrdI* cuts Ad5 at positions 5458, 7039, 14004, 15593, 17257, and 21023. The 21023-35938 bp fragment then can be used as the right arm and purified from the other fragments by gel electrophoresis. These two fragments are then transfected with the *PmeI/BamHI* fragment from pACT H11(RK32) into 293 cells. *PmeI* cuts at position 13258 in Ad5, and *BamHI* cuts at position 21562 in Ad5. Similar techniques can be employed to produce AdZ.H(RKKK),.

Example 16

This example describes the construction of vectors containing a short-shafted fiber protein.

The method described herein for construction of an adenovirus from the transfer plasmid p193 F5F9Kshort can similarly be employed for the construction of other adenoviral vectors from other short-shafted fibers. Namely, the transfer plasmid p193 F5F9Kshort, which contains the essential E4 region of adenovirus, was cut with *SalI* and transfected into 293 cells, which had been infected one hour earlier with the adenoviral vector AdSE.E4Gus. AdSE.E4Gus lacks the E4 region of the adenoviral genome, and cannot replicate in 293 cells in the absence of complementation for the E4 genes. Thus, only when the AdSE.E4Gus DNA recombines with the p193 F5F9K short plasmid DNA to obtain the E4 genes is the vector able to replicate in 293 cells. During this recombination event, the newly formed vector also acquires the mutated fiber protein coding sequences

encoded by the plasmids. Viable recombinant E4⁺ adenovirus containing the F5F9Kshort fiber chimera were then isolated by plaquing the transfected cell lysates 5 days after transfection. The resultant vector

5 AdSE.F5F9Kshort was isolated and purified by standard virological techniques involving two successive rounds of plaquing on 293 cells. The vector was verified to contain the correct insert by PCR and restriction analysis of viral DNA. Oligonucleotide primers, which

10 prime on either side of the fiber gene, confirmed that the PCR product was of the correct size for that encoded by a shortened chimeric fiber gene. Restriction analysis of the vector DNA showed that the new vector contained the correct restriction sites that are unique to the Ad9

15 fiber knob.

Example 17

This example describes the construction of adenoviral vectors that contain short-shafted fiber

20 proteins and chimeric penton base proteins incorporating UTV or UTV-like sequences.

For this construction, AdS.F9sK viral DNA can be digested with *Xmn*I, as described above. The plasmid pACT H11(RK32) is then cut with the restriction enzymes *Pme*I

25 and *Bam*HI. The restriction digested viral and plasmid DNAs are purified and transfected into 293 cells. The resultant vectors are isolated by two successive rounds of plaque purification on 293 cells, and are verified to contain the correct sequence in the penton base region by

30 restriction analysis of viral DNA obtained from vector-infected 293 cells by Hirt extraction.

Sequencing of PCR products generated by amplifying the region of the insert from the viral DNA also can be used to verify the presence of the insert. Western

35 analysis of the chimeric penton base protein on a polyacrylamide gel should show a corresponding shift in the mobility of the chimeric penton base compared to

unmodified penton base that is consistent with the presence of additional nonnative amino acid sequences (and absence of native amino acid sequences).

Other plasmids for which maps are presented herein and which were not made into an adenoviral vector can be made by utilizing the same or a slightly modified version of the procedure outlined above. In particular, the short-shafted fiber protein can be incorporated into an adenovirus having a "spiked" chimeric penton base protein that furthermore optionally can incorporate a UTV or UTV-like sequence.

Example 18

This example describes the construction of adenoviral vectors that contain short-shafted fiber proteins and chimeric hexon proteins incorporating UTV or UTV-like sequences.

Viral DNA can be isolated from a short-shafted vector such as AdZ.F9sK and cut with the restriction enzymes described above for making vectors comprising UTV-containing chimeric hexon proteins. All other steps are the same as described, for instance, in Example 17. The resultant vector should contain the short-shafted fiber protein and the chimeric hexon protein incorporating UTV or UTV-like sequences. Moreover, this approach can be employed with a variety of transfer plasmids comprising different chimeric hexon proteins. In particular, the short-shafted fiber protein can be incorporated into an adenovirus having a "spiked" chimeric hexon protein that furthermore optionally can incorporate a UTV or UTV-like sequence.

Example 19

This example describes an evaluation of vectors, particularly adenoviral vectors, according to the invention, which comprise UTV or UTV-like sequences.

For instance, to confirm that the addition of UTV or UTV-like sequences has no effect on virus assembly, the virus growth kinetics of the vectors can be assessed. As representative of a UTV-sequence containing plasmid, the growth behaviour of pAd.F(pK7) was monitored and compared to that of wild-type adenovirus (Ad5), as well as the adenoviral vector AdZ.F(RGD), which contains an insertion of a RGD peptide motif present in the sequence SACDCRGDCFCGTS [SEQ ID NO:93]. For these studies, 293 cells were infected at a multiplicity of infection of 5 active virus particles/cell with either Ad5, AdZ.F(RGD) or AdZ.F(pK7), and the number of infectious particles (fluorescent focus units (FFU)) produced per cell was determined following the harvesting of the cells at 1, 2, and 3 days post infection. The titers of AdZ.F(RGD) or AdZ.F(pK7) were somewhat lower, but not dramatically different than the titer of Ad5. As can be seen in Figure 28, the peak titers of AdZ.F(RGD) and AdZ.F(pK7) were 80% and 56%, respectively, that of Ad5. These results confirm that the growth kinetics of the two vectors are not substantially affected by the addition of sequences, particularly a UTV or UTV-like sequence, onto the end of the fiber protein. The results also suggest that further vectors comprising UTV or UTV-like sequences will not exhibit aberrant growth behaviour.

Furthermore, vectors containing UTV or UTV-like sequences can be evaluated for their ability to bind to cells or deliver genes to be inhibited by negatively-charged molecules (e.g., heparin, heparan sulfate, chondroitin sulfate, etc.), soluble adenoviral coat proteins, or by pretreatment of cells with agents (e.g., chondroitinase, heparinase, sialidase, etc.) that cleave such negatively-charged moieties (see, e.g., Wickham et al., Nature Biotechnology, 14, 1570-1573 (1996), as well as the preceding Examples). Soluble fiber protein will not inhibit the majority of the binding of a UTV vector to a cell (Wickham et al. (1996), *supra*). These results

suggest that the incorporation of UTV or UTV-like sequences into penton, hexon or fiber will not impair the ability of a recombinant adenovirus containing the chimeric coat protein to effect gene delivery.

5 Also, studies of infection *in vivo*, or *in vitro* transfections done in the presence of whole blood, can be employed to confirm that the UTV vectors of the present invention are not limited for systemic delivery due to saturation of the polycations on the recombinant
10 adenoviruses with polyanions in the blood. In the event that such binding impedes the capability of a particular virus for target cell transduction, the virus can be administered in a higher dose, preferably with provision being made to reduce any immune response associated with
15 such a higher dose (e.g., administration of another serotype of adenoviral vector, or techniques described in PCT International Application WO 96/12406; Mastrangeli et al., Human Gene Therapy, 7, 79-87 (1996)).

20 All of the references cited herein, including patents, patent applications, and publications, are hereby incorporated in their entireties by reference to the same extent as if each reference were set forth in its entirety herein.

25 While this invention has been described with an emphasis upon preferred embodiments, it will be apparent to those of ordinary skill in the art that variations in the preferred embodiments can be prepared and used and that the invention can be practiced otherwise than as
30 specifically described herein. The present invention is intended to include such variations and alternative practices. Accordingly, this invention includes all modifications encompassed within the spirit and scope of the invention as defined by the following claims.

35

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: GenVec, Inc.
- (ii) TITLE OF INVENTION: VECTORS AND METHODS FOR GENE TRANSFER TO CELLS
- (iii) NUMBER OF SEQUENCES: 12
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Leydig, Voit & Mayer, Ltd.
 - (B) STREET: Two Prudential Plaza, Suite 4900
 - (C) CITY: Chicago
 - (D) STATE: IL
 - (E) COUNTRY: USA
 - (F) ZIP: 60601
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: PCT
 - (B) FILING DATE: 27-NOV-1996
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/563,368
 - (B) FILING DATE: 28-NOV-1995
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (C) REFERENCE/DOCKET NUMBER: 74862

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 8 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Lys Lys Lys Lys Lys Lys Lys Lys
1 5

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 8 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Arg Arg Arg Arg Arg Arg Arg
1 5

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(C) OTHER INFORMATION: Xaa is Lys or Arg

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Gly Ser Asn Lys Glu Ser Phe Val Leu Lys Lys Lys Lys Lys Lys
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Ala Gly Ser Asn Lys Asn Lys Glu Ser Phe Val Leu Lys Lys Lys Lys
1 5 10 15
Lys Lys

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 base pairs
(B) TYPE: nucleic acid

81

- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGA TCC AA
Gly Ser
1

8

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGA TCC AAT AAA GAA TCG TTT GTG TTA TGT
Gly Ser Asn Lys Glu Ser Phe Val Leu Cys
1 5 10

30

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GCC GGA TCC AAC AAG AAT AAA GAA TCG TTT GTG TTA
Ala Gly Ser Asn Lys Asn Lys Glu Ser Phe Val Leu
1 5 10

36

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 55 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TATGGAGGAT CCAATAAAGA ATCGTTTGTG TTATGTTTCA ACGTGTTCAT TTTTC

55

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 57 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AATTGAAAAA TAAACACGTT GAAACATAAC ACAAACGATT CTTTATTGGA TCCTCCA 57

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 43 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TCCCCCGGG TCTAGATTAG GATCCTTCTT GGGCAATGTA TGA 43

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CGTGTATCCA TATGACACAG A 21

WHAT IS CLAIMED IS:

1. A chimeric adenovirus coat protein comprising a nonnative amino acid sequence, wherein said chimeric
5 adenovirus coat protein is able to direct entry into cells of a vector comprising said chimeric adenovirus coat protein that is more efficient than entry into cells of a vector that is identical except for comprising a wild-type adenovirus coat protein rather than said
10 chimeric adenovirus protein.

2. The chimeric adenovirus coat protein of claim 1, wherein said chimeric adenovirus coat protein binds a novel binding site present on the cell surface.

15

3. The chimeric adenovirus coat protein of claim 1 or 2, wherein said nonnative amino acid sequence is inserted into or in place of an internal coat protein sequence.

20

4. The chimeric adenovirus coat protein of claim 1 or 2, wherein said nonnative amino acid sequence is at or near the C-terminus of said protein.

25

5. The chimeric adenovirus coat protein of claim 1 or 2, wherein said nonnative amino acid sequence is in an exposed loop of said protein.

6. The chimeric adenovirus coat protein of any of
30 claims 1-5, wherein said nonnative amino acid sequence comprises a UTV or UTV-like sequence which is linked to said protein by a spacer sequence of from about 3 to about 400 amino acids.

35

7. The chimeric adenovirus coat protein of claim 6, wherein said spacer sequence is selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID

NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:20, SEQ ID
NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID
NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID
NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:46, SEQ ID
5 NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID
NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID
NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:73, SEQ ID
NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:90, and SEQ
ID NO:93.

10

8. The chimeric adenovirus coat protein of claim 6 or 7, wherein said UTV or UTV-like sequence comprises from about 3 to about 30 amino acids.

15

9. The chimeric adenovirus coat protein of any of claims 6-8, wherein said UTV or UTV-like sequence comprises amino acids selected from the group consisting of lysine, arginine, and histidine.

20

10. The chimeric adenovirus coat protein of any of claims 6-8, wherein said UTV or UTV-like sequence comprises amino acids selected from the group consisting of aspartate and glutamate.

25

11. The chimeric adenovirus coat protein of any of claims 6-8, wherein said UTV or UTV-like sequence is selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:90, and SEQ ID NO:93, and wherein said sequence may be deleted at the C-terminus by 1, 2, 3, 4, or 5 residues.

30

35

12. The chimeric adenovirus coat protein of any of claims 6-8, wherein said UTV or UTV-like sequence is selected from the group consisting of SEQ ID NO:4 and SEQ ID NO:5, and wherein the length of the polylysine string can vary from about 3 to about 30 residues.

13. The chimeric adenovirus coat protein of any of claims 1-12, wherein said chimeric adenovirus coat protein is a fiber protein.

14. The chimeric adenovirus coat protein of any of claims 1-12, wherein said chimeric adenovirus coat protein is a penton base protein.

15. 15. The chimeric adenovirus coat protein of any of claims 1-12, wherein said chimeric adenovirus coat protein is a hexon protein.

20 16. An isolated and purified nucleic acid encoding a chimeric adenovirus coat protein of any of claims 1 to 15, wherein said nucleic acid comprises a sequence selected from the group consisting of SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:72, SEQ ID NO:75, SEQ ID NO:77, and SEQ ID NO:89.

30 17. The nucleic acid of claim 16, wherein said chimeric adenovirus coat protein is a fiber protein.

35 18. The nucleic acid of claim 16, wherein said chimeric adenovirus coat protein is a penton base protein.

19. The nucleic acid of claim 16, wherein said chimeric adenovirus coat protein is a hexon protein.

20. A vector that comprises a chimeric adenovirus coat protein of any of claims 1-15.

21. The vector of claim 20, wherein said vector is a viral vector selected from the group consisting of nonenveloped viruses.

22. The vector of claim 20, wherein said vector is an adenoviral vector.

23. The vector of any of claims 20-22, wherein said vector further comprises a nucleic acid comprising a passenger gene.

24. The vector of claim 23, wherein said nucleic acid is inserted into the adenoviral genome.

25. The vector of claim 23, wherein said nucleic acid is attached to said chimeric adenovirus coat protein by means of a protein/DNA interaction.

26. A vector selected from the group consisting of GV10 UTV, AdZ.H(RKKK)₂, AdZ.H(RKKK)₃, AdZ.F(RKKK)₂, AdZ.F(RKKK)₃, AdZ.F(pK7), AdZ.F(pGS), AdZ.F(tat), AdZ.F52K(RKKK)₂, and AdZ.F52K(RKKK)₃.

27. The vector of any of claims 20-25, wherein said chimeric adenovirus coat protein is a fiber protein.

28. The vector of any of claims 20-25, wherein said chimeric adenovirus coat protein is a penton base or hexon protein.

29. The vector of claim 28, wherein said vector does not comprise wild-type fiber protein.

30. The vector of claim 28, wherein said vector
5 comprises a short-shafted fiber protein.

31. A transfer vector that encodes the chimeric adenovirus coat protein of any of claims 1-15.

10 32. The transfer vector of claim 31, wherein said transfer vector is selected from the group consisting of nonenveloped viruses.

33. The transfer vector of claim 31, wherein said
15 transfer vector is an adenoviral vector.

34. The transfer vector of any of claims 31-33,
wherein said transfer vector further comprises a nucleic
acid encoding a passenger gene.

20

35. A transfer vector selected from the group
consisting of p193 (F5*), pAd BS 59-100 UTV, p193NS (F5*)
pGS(pk7), p193NS (F5*) pGS(null), pBSS 75-100 pGS (null),
pBSS 75-100 pGS(RK32), pBSS 75-100 pGS(RK33), pBSS 75-100
25 pGS(tat), p193 F5F2K, p193 F5F2K(RKKK)₂, p193
F5F2K(RKKK)₃, pACT (RKKK)₂, pACT (RKKK)₃, pACT H11, pACT
H11 (RKKK)₂, pACT H11 (RKKK)₃, pACT H12, pACT H12 (RKKK)₂,
pACT H12 (RKKK)₃, pSP2alpha, pSP2alpha2, pSP2alpha2(RKKK)₂,
pSP2alpha2(RKKK)₃, pACT2alpha2 (RKKK)₂, pACT2alpha2
30 (RKKK)₃, pACT H11 2alpha2 (RKKK)₂, pACT H11 2alpha2
(RKKK)₃, pACT H12 2alpha2 (RKKK)₂, and pACT H12 2alpha2
(RKKK)₃.

36. The transfer vector of any of claims 31-34,
35 wherein said coat protein is a fiber protein.

37. The transfer vector of any of claims 31-34, wherein said coat protein is a penton base or hexon protein.

5 38. A method of increasing the efficiency of entry into cells of a vector comprising a coat protein, which comprises replacing a coat protein of said vector with the chimeric adenovirus coat protein of any of claims 1-15.

10

39. A method of genetically modifying a cell which comprises contacting said cell with a vector of any of claims 20-37.

15

40. A method of genetically modifying a cell which comprises:

(a) obtaining a vector comprising the coat protein of any of claims 1-15,

20

(b) contacting said vector with a nucleic acid encoding a passenger gene such that said chimeric protein and said nucleic acid bind due to a protein/DNA interaction, and

(c) contacting said cell with said vector comprising bound nucleic acid.

25

41. A host cell comprising a vector of any of claims 20-37.

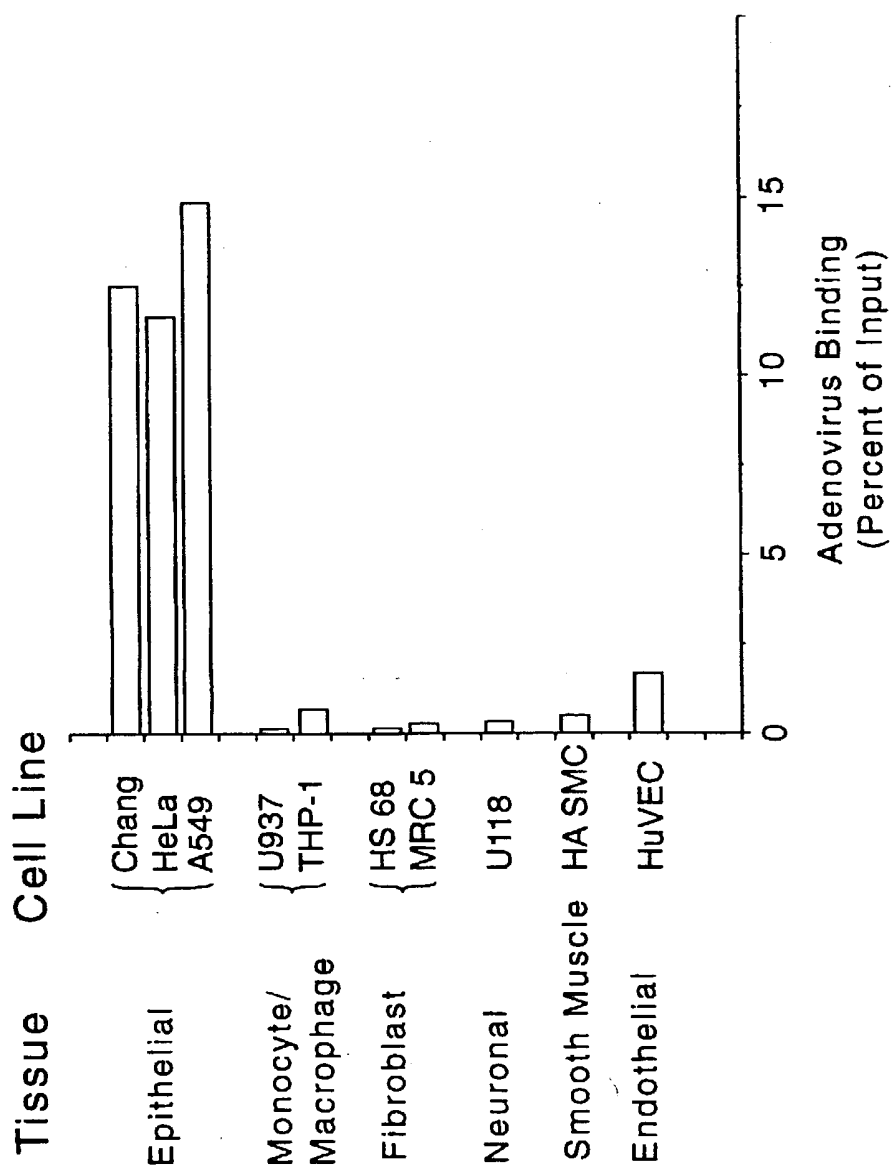
30 42. A method of making the chimeric adenovirus coat protein of claim 1, which method comprises introducing a spacer sequence into the coding sequence of said chimeric adenovirus coat protein prior to any stop codon or polyadenylation signal.

35

43. The method of claim 42, wherein said spacer sequence comprises the sequence of SEQ ID NO:6 or a conservatively modified variant thereof.

44. A method of making the vector GV10 UTV, wherein said method comprises providing to a cell that does not comprise any E4 complementing sequences: (a) a linear
5 vector comprising the chimeric fiber and a wild-type E4 gene, and (b) a linear vector that is E4⁻.

1/30

**FIG. 1**

2/30

Ala Gln Glu *
GCC CAA GAA TAA AGA ATC GTT TGT GTT ATG TTT CAA CGT [SEQ ID NO:13]

TRANSCRIPTION

GCC CAA GAA UAA AGA AUC GUU UGU GUU AAA AAA AAA AAA AAA AAA [SEQ ID NO:14]

TRANSLATION

Ala Gln Glu *

FIG. 2A

3/30

Ala Gln Glu Gly Ser Asn Lys Glu Ser Phe Val Leu Cys [SEQ ID NO:16]

GCC CAA GAA GGA TCC AAT AAA GAA TCG TTT GTG TTA TGT [SEQ ID NO:15]

BamHI PolyA

TRANSCRIPTION

GCC CAA GAA GGA UCC AAU AAA GAA UCG UUU GUG UUA AAA AAA AAA AAA...[SEQ ID NO:17]

TRANSLATION

Ala Gln Glu Gly Ser Asn Lys Glu Ser Phe Val Leu Lys Lys Lys Lys...[SEQ ID NO:18]

FIG. 2B

4/30

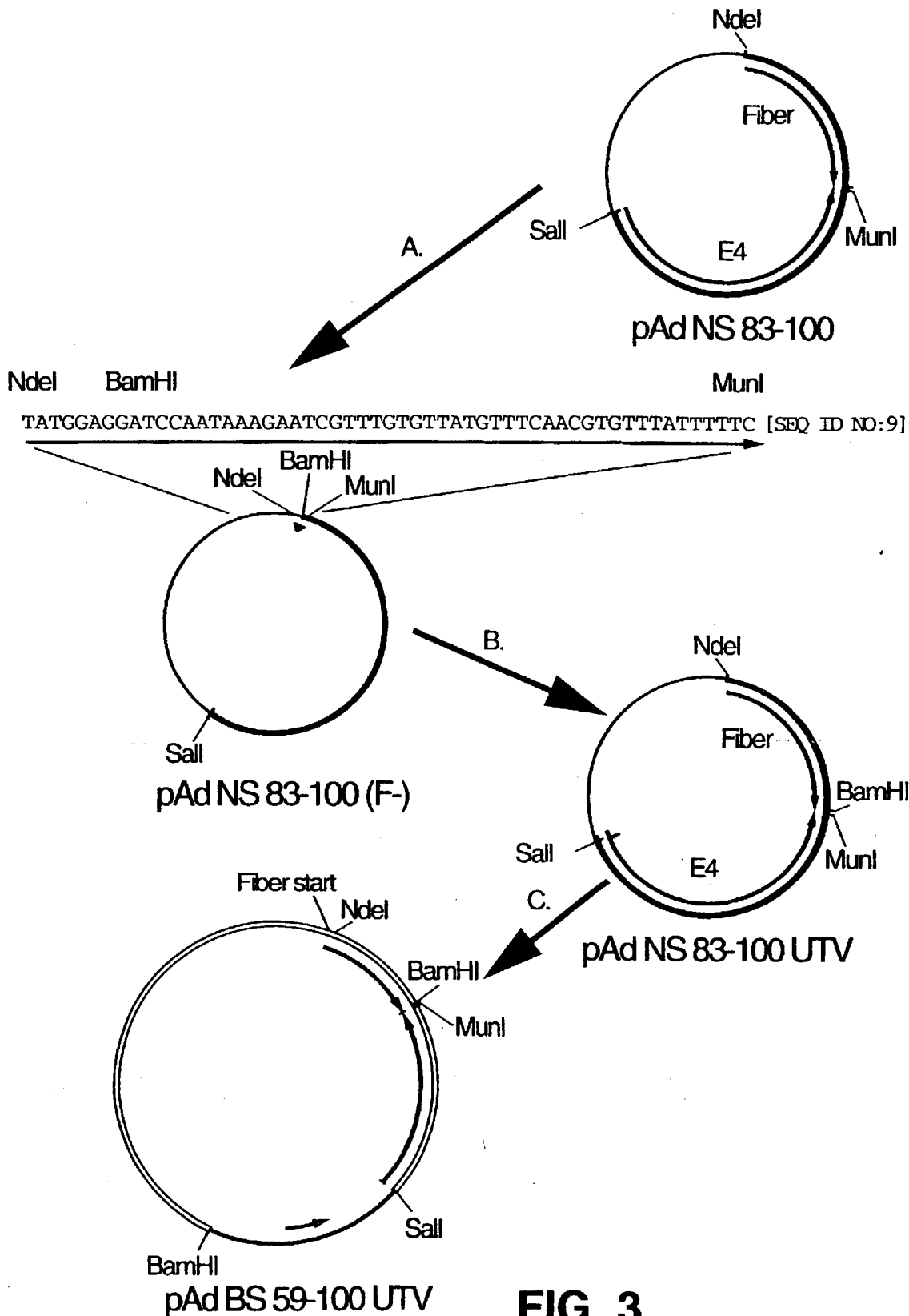


FIG. 3

5/30

TAT GGA GGA TCC AAT AAA GAA TCG TTT GTG TTA TGT TTC AAC GTG TTT AAT TTT C [SEQ ID NO:9]
 NdeI BamHI PolyA
 polyA MunI

FIG. 4A

AAT TGA AAA ATA AAC ACG TTG AAA CAT AAC ACA AAC GAT TCT TTA TTG GAT CCT CCA [SEQ ID NO:10]
 MunI PolyA
 BamHI NdeI

FIG. 4B

TCCC CCGGGG TCTAGA TTA GGA TCC TCC TTG GGC AAT GTA TGA [SEQ ID NO:11]
 XmaI XbaI BamHI

FIG. 4C

CGT GTA TCC ATA TGA CAC AGA [SEQ ID NO:12]
 NdeI

FIG. 4D

6/30

WT UTV

62 kD → — —

FIG. 5

7/30

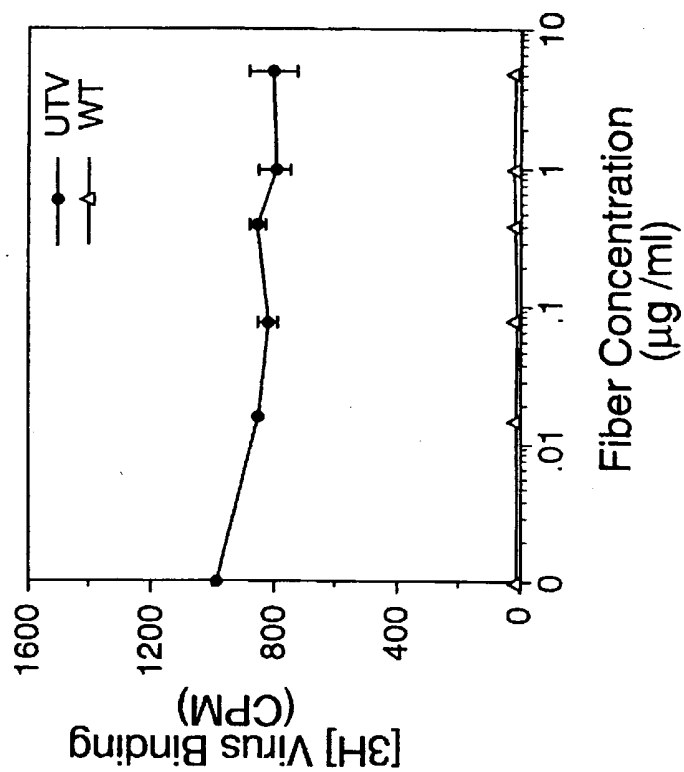


FIG. 6B

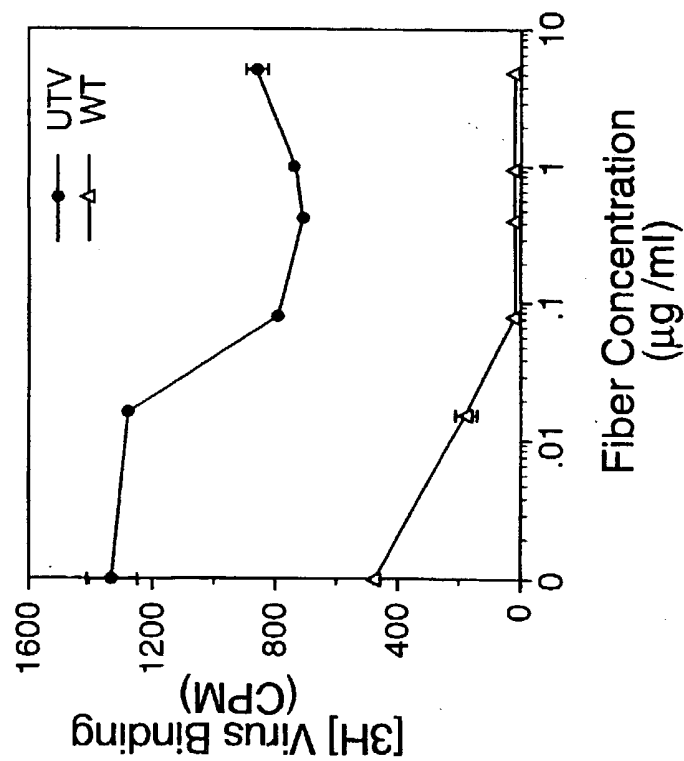
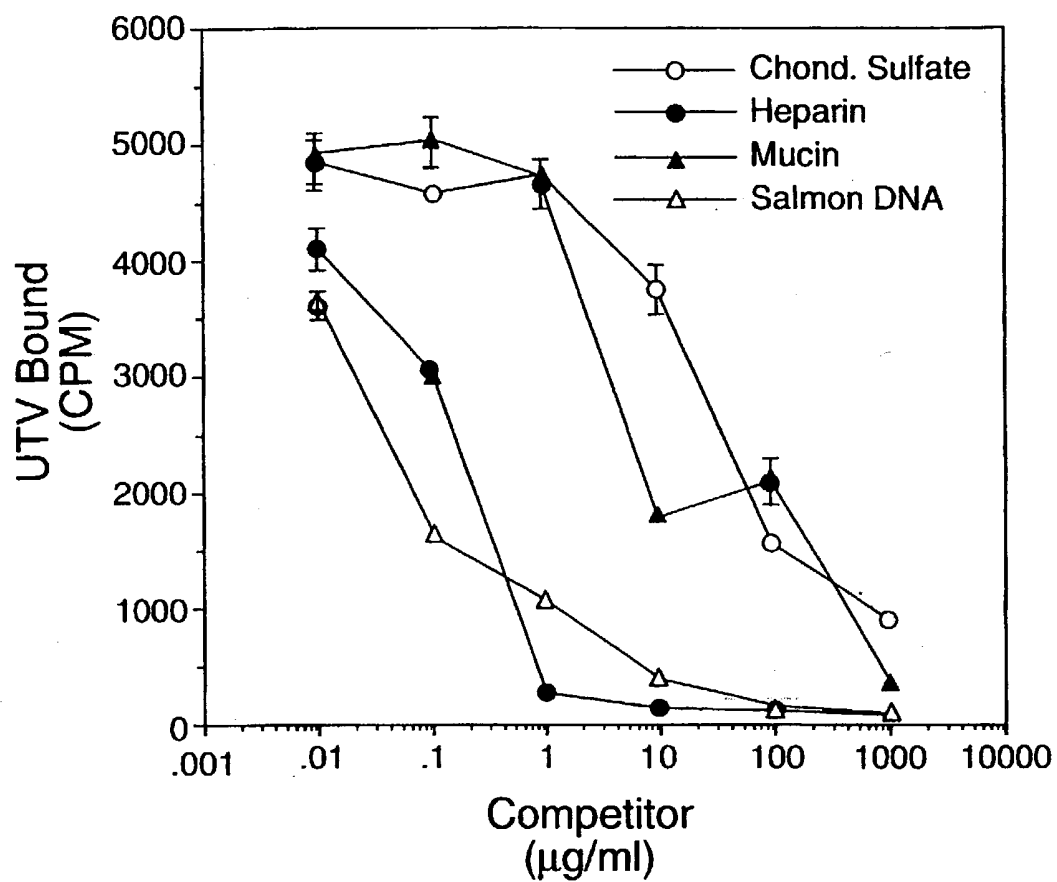
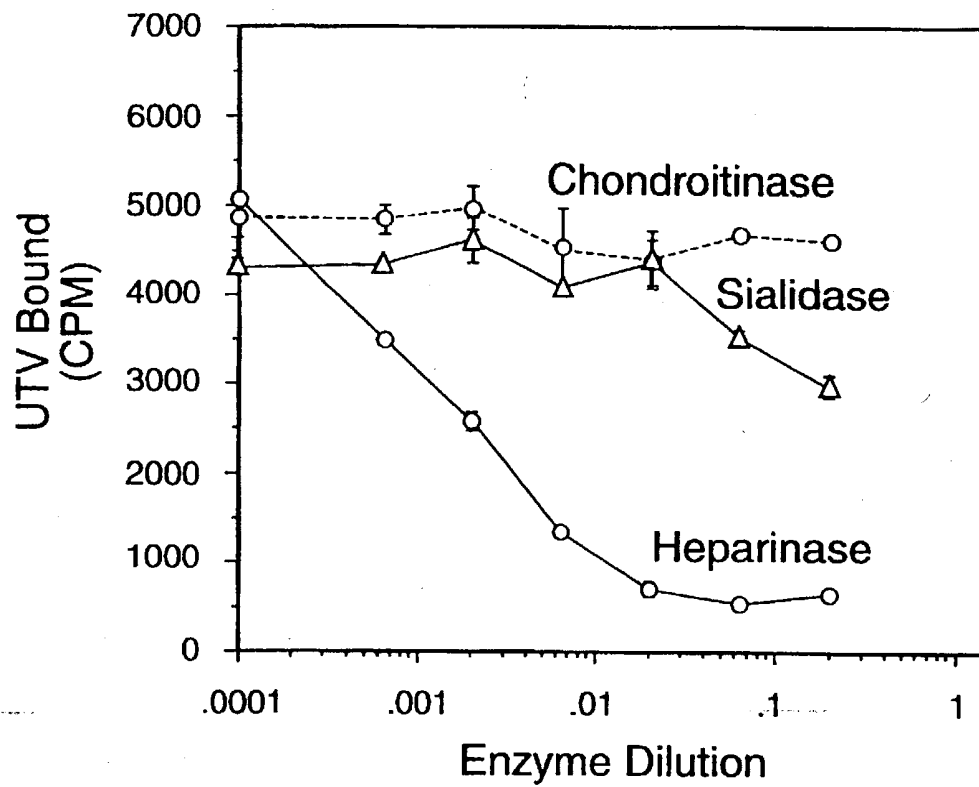


FIG. 6A

8/30

**FIG. 7**

9/30

**FIG. 8**

10/30

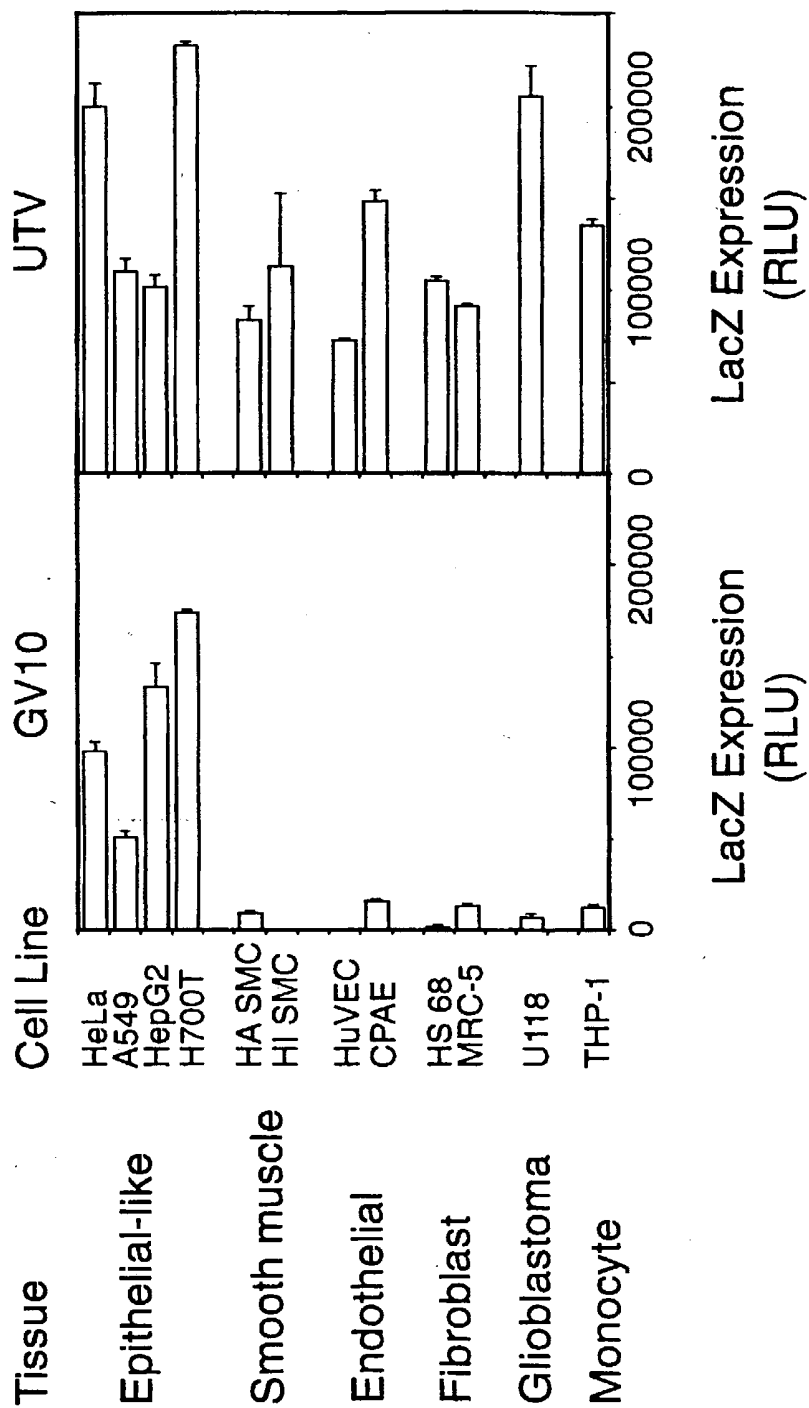
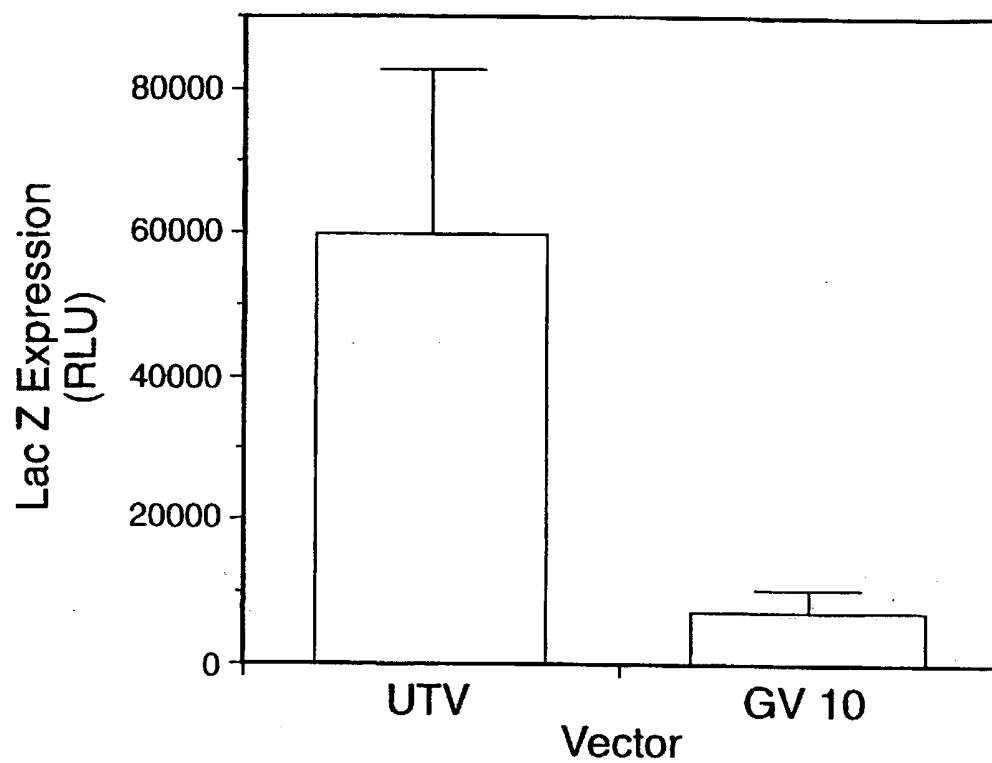
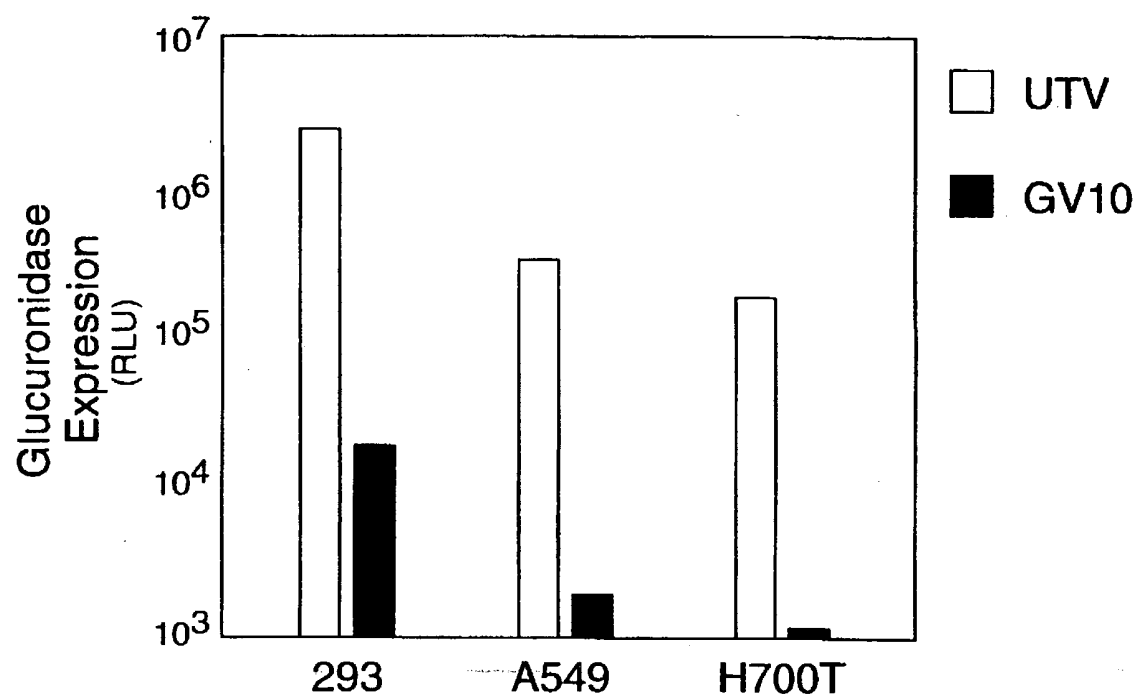


FIG. 9

11/30

**FIG. 10**

12/30

**FIG. 11**

13/30

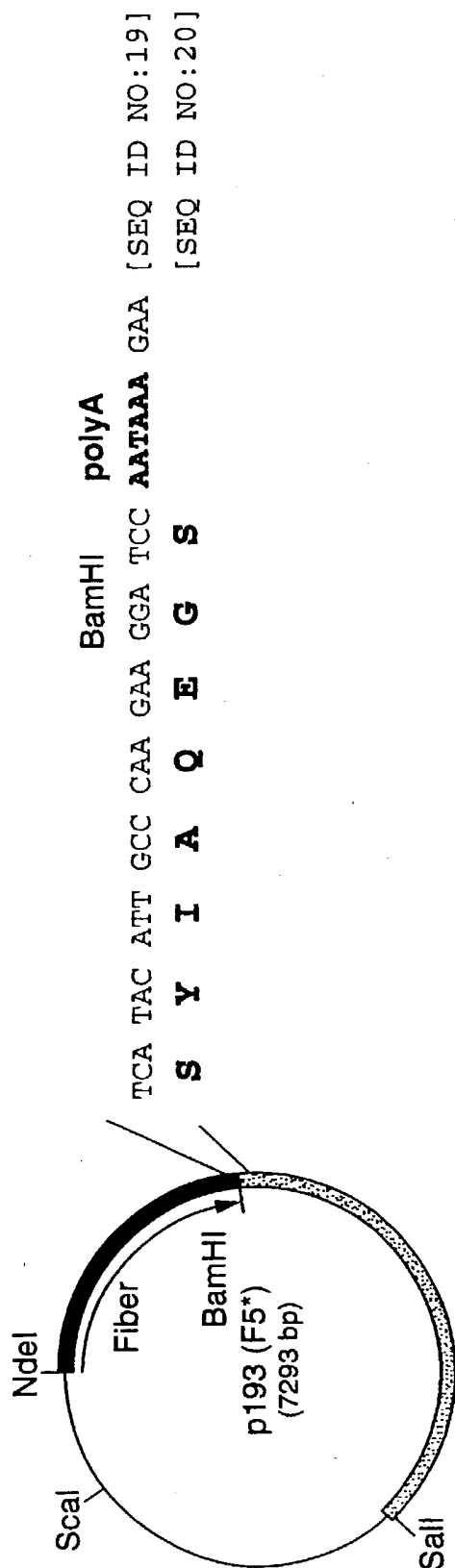
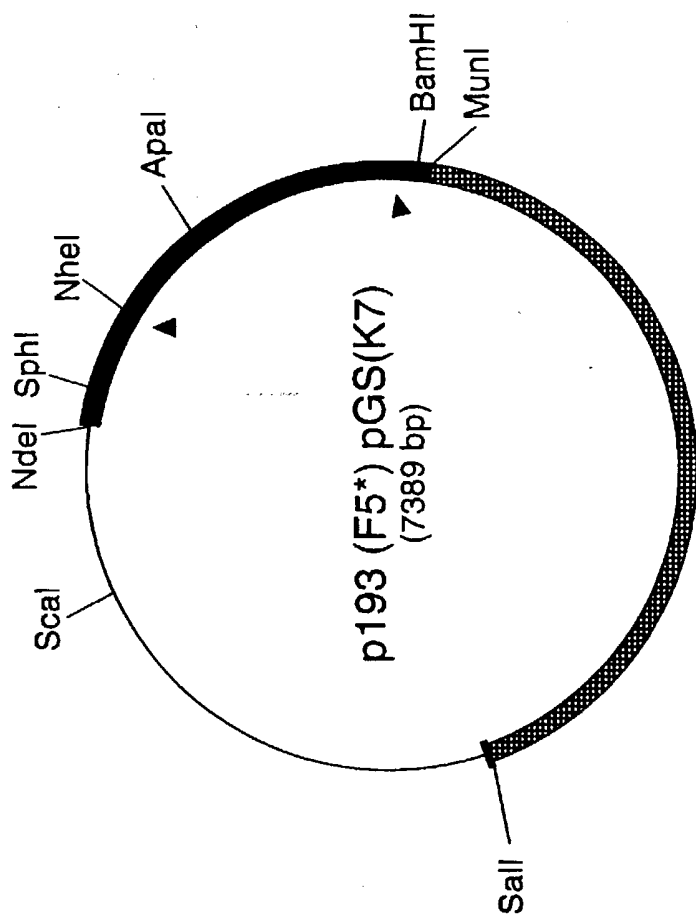


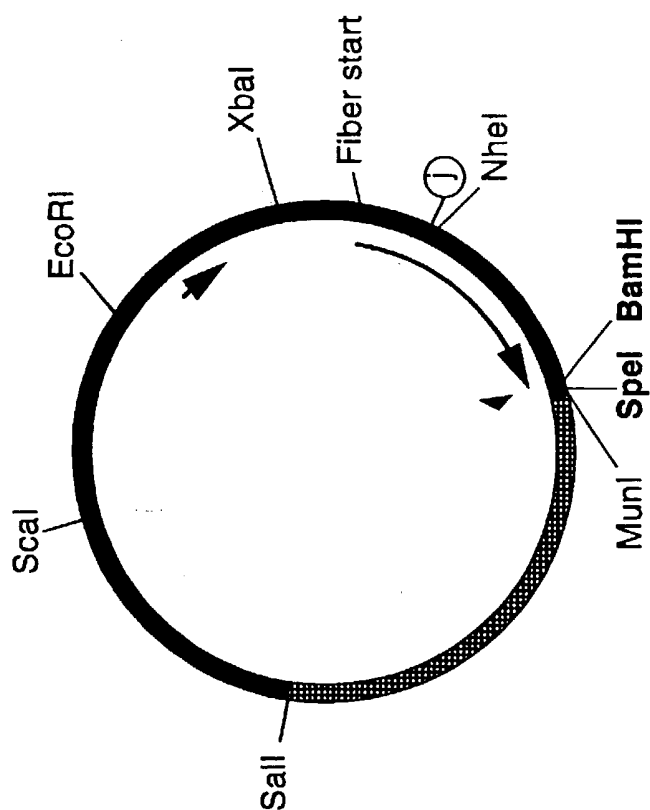
FIG. 12



Accession	Gene	Protein	Seq ID	Seq ID
GGATCAGGATCAGGTTCAGGAGTGGCTCTAAAGAAGAAAAAGAAAGTAA			[SEQ ID NO:21]	[SEQ ID NO:21]
GlySerGlySerGlySerGlySerGlySerLysLysLysLysLys			[SEQ ID NO:22]	[SEQ ID NO:22]

FIG. 13

pBSS75-100ΔE3 pGS (null)
(9972)

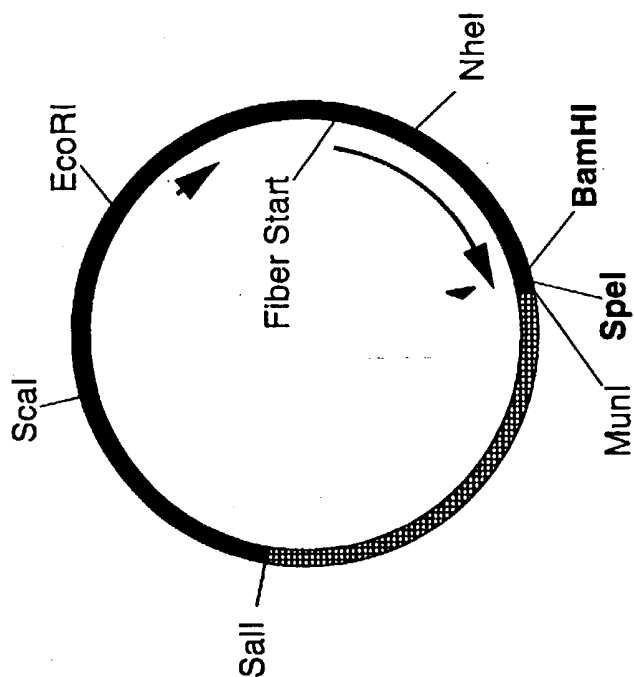


BamHI
GCCCCAAGAAGGATCCCGGTTTCAGGATCTGGCAGTGGCTCGACTAGTTAA [SEQ ID NO:23]
AlaGlnGluGlySerGlySerGlySerGlySerGlySerGlySerGlySerThrSer [SEQ ID NO:24]

FIG. 14

16/30

pBSS75-100ΔE3 pGS(RKKK)2
(9999bp)



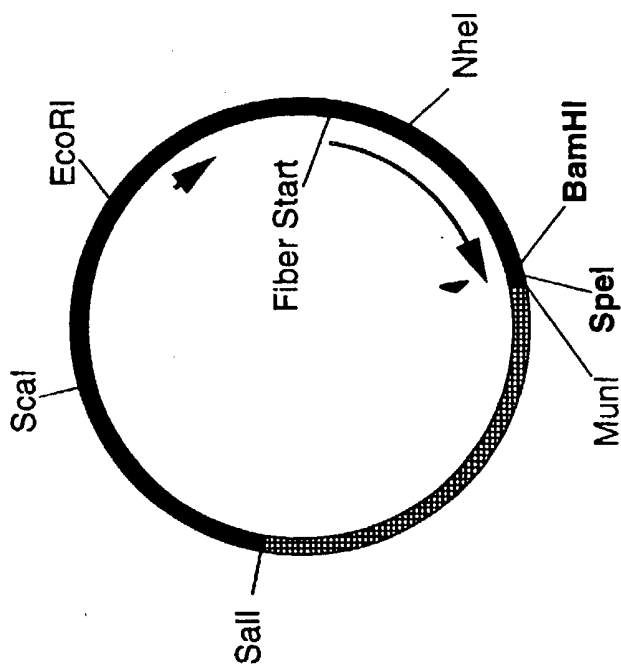
BamHI

GCCCAAGAAGGATCCGGTTCAGGATCTGGCAGTGGCTCGACTAGAAAGAAACGCAAAAAGAAAGACTAGTTAA [SEQ ID NO:25]
AlaGlnGluGlySerGlySerGlySerGlySerGlySerThrArgLysLysLysArgLysLysLysLysThrSer [SEQ ID NO:26]

SpeI

FIG. 15

pBSS75-100ΔE3 pGS(RKKK)3
(10010bp)



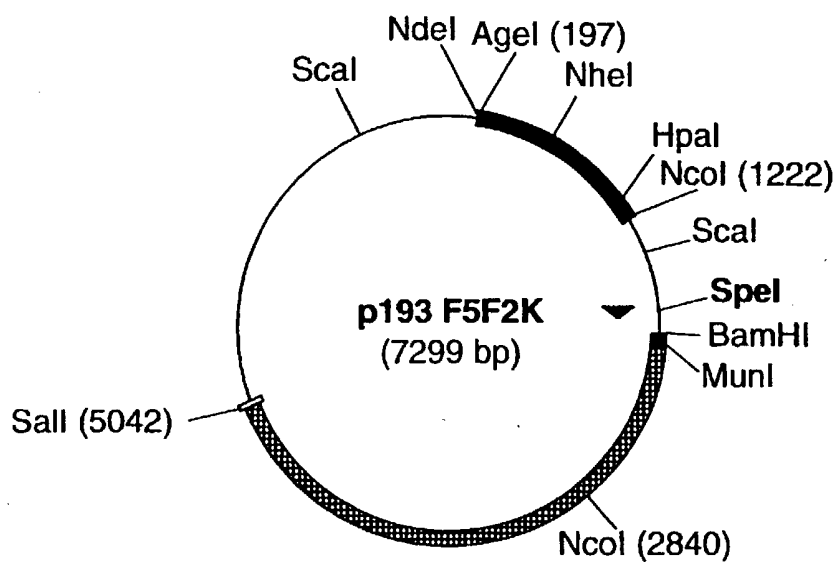
BamHI

Spel

GCCCCAAGAGGATCCGGTTCAGGATCTGGCAGTGGCTCGACTAGAAAGAAGCGCAAAAAAAGAAAGAACTAGTTAA [SEQ ID NO:27]
 AAlaGlnGluGlySerGlySerGlySerGlySerGlySerThrArgLysLysLysArgLysLysLysArgLysLysThrSer [SEQ ID NO:28]

FIG. 16

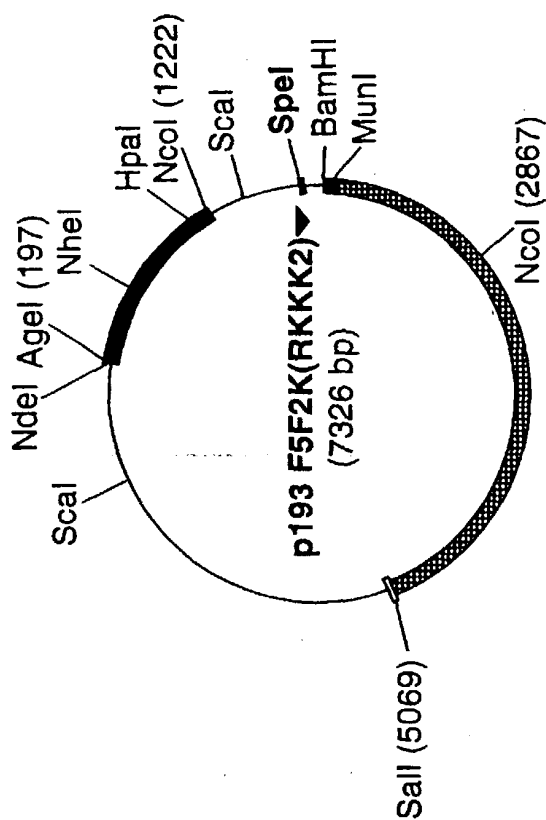
18/30

**Spel**ATTACACTTAATGGC**ACTAGT**GAATCCACAGAACT [SEQ ID NO:29]

IleThrLeuAsnGlyThrSerGluSerThrGluThr [SEQ ID NO:30]

FIG. 17

19/30

**Spel**

ATTACACTTAATGCCACTAGAAAGAGAAACGCAAAAGAGACTAGTGAATCCACAGAAACT [SEQ ID NO:31]
IleThrLeuAsnGlyThrArgLysLysLysArgLysLysLysThrSerGluSerThrGluThr [SEQ ID NO:32]

FIG. 18

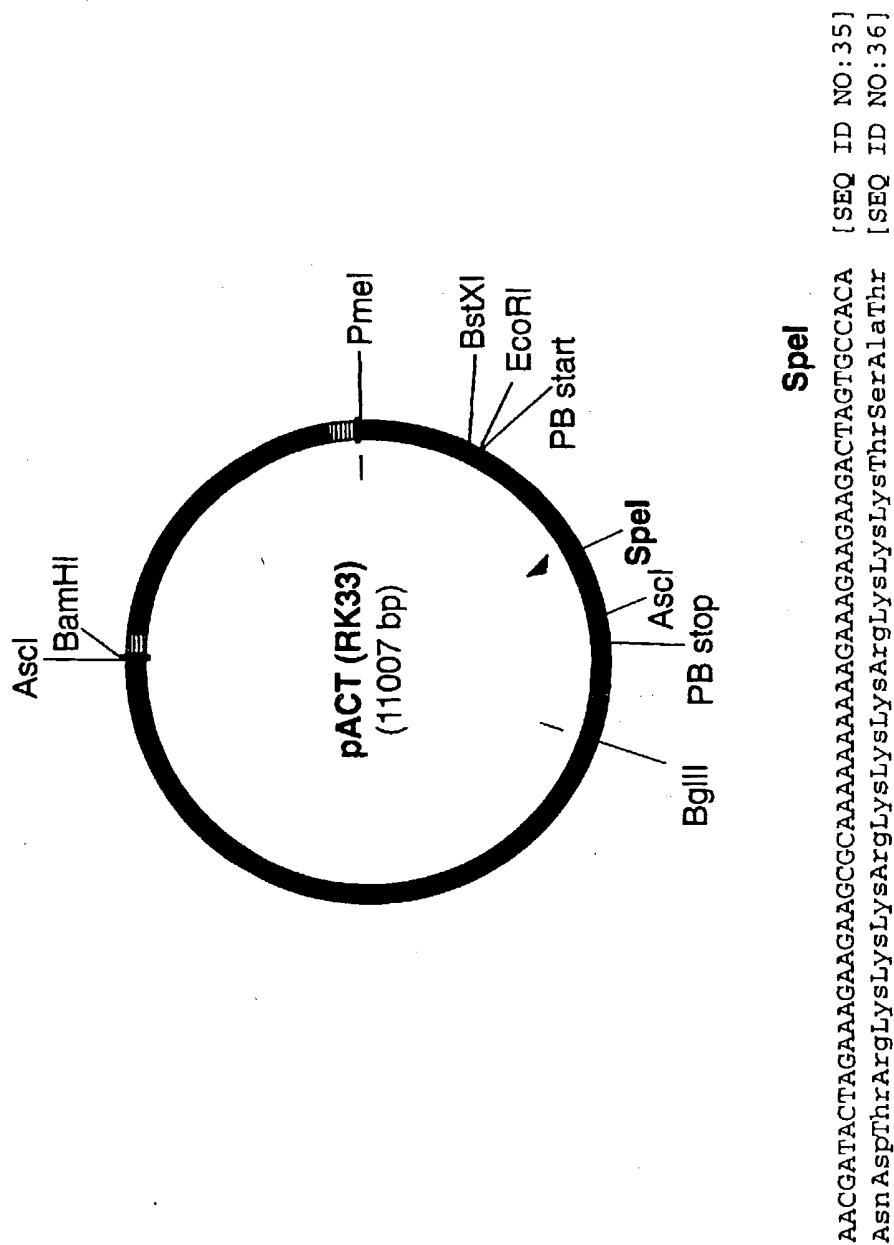
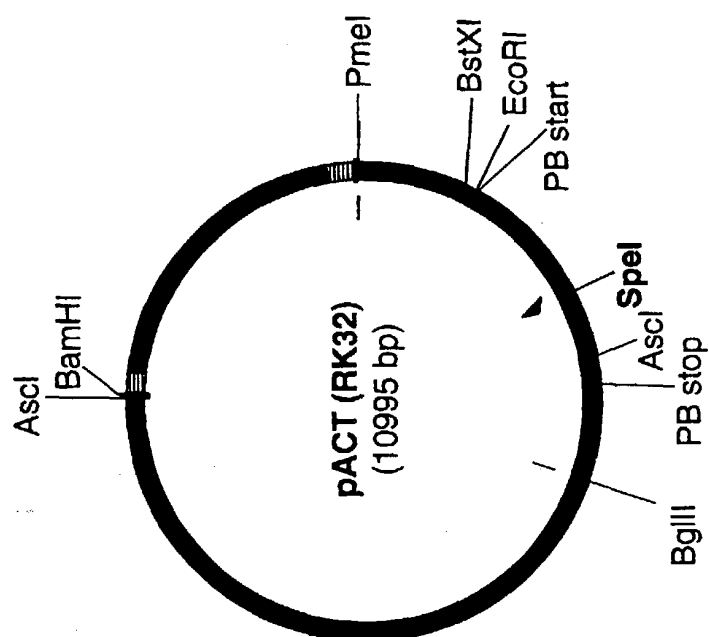
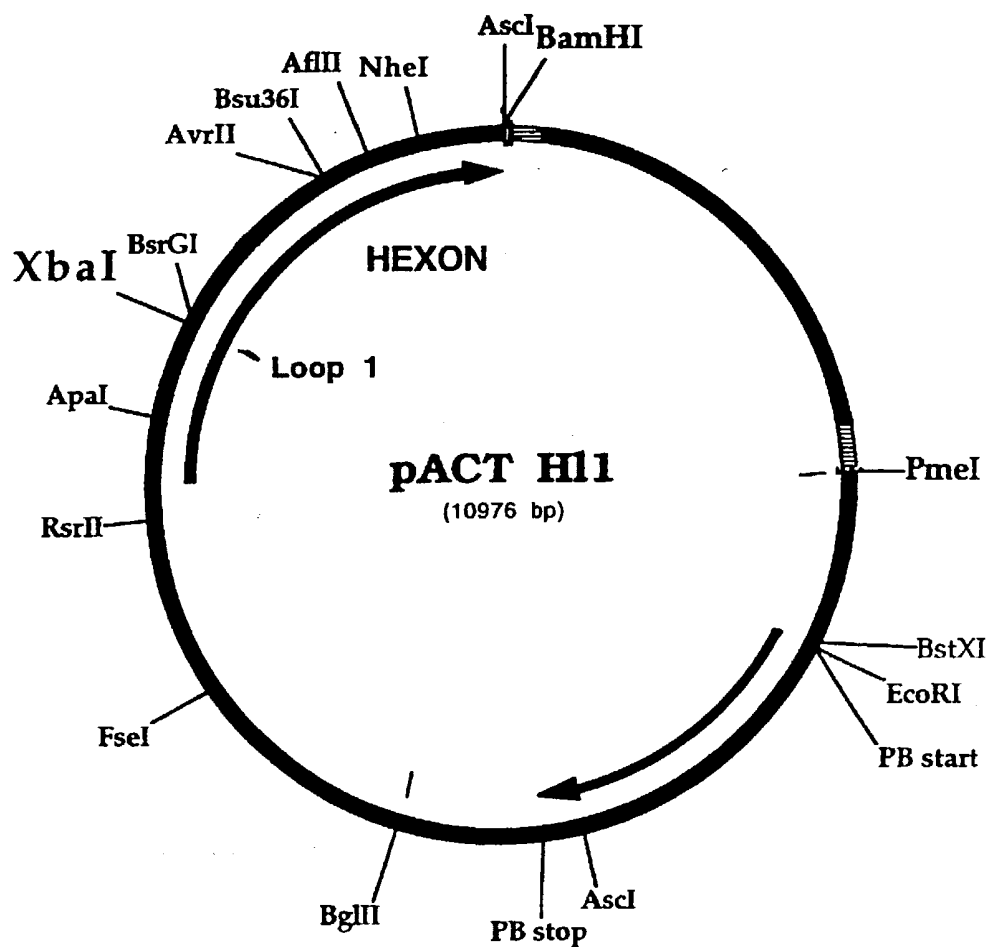


FIG. 20

**Spel**

AACGATACTAGAAACAAGAGAGAAAGAAGACTAGTGCCACA [SEQ ID NO:37]
AsnAspThrArgLysLysLysArgLysLysLysThrSerAlaThr [SEQ ID NO:38]

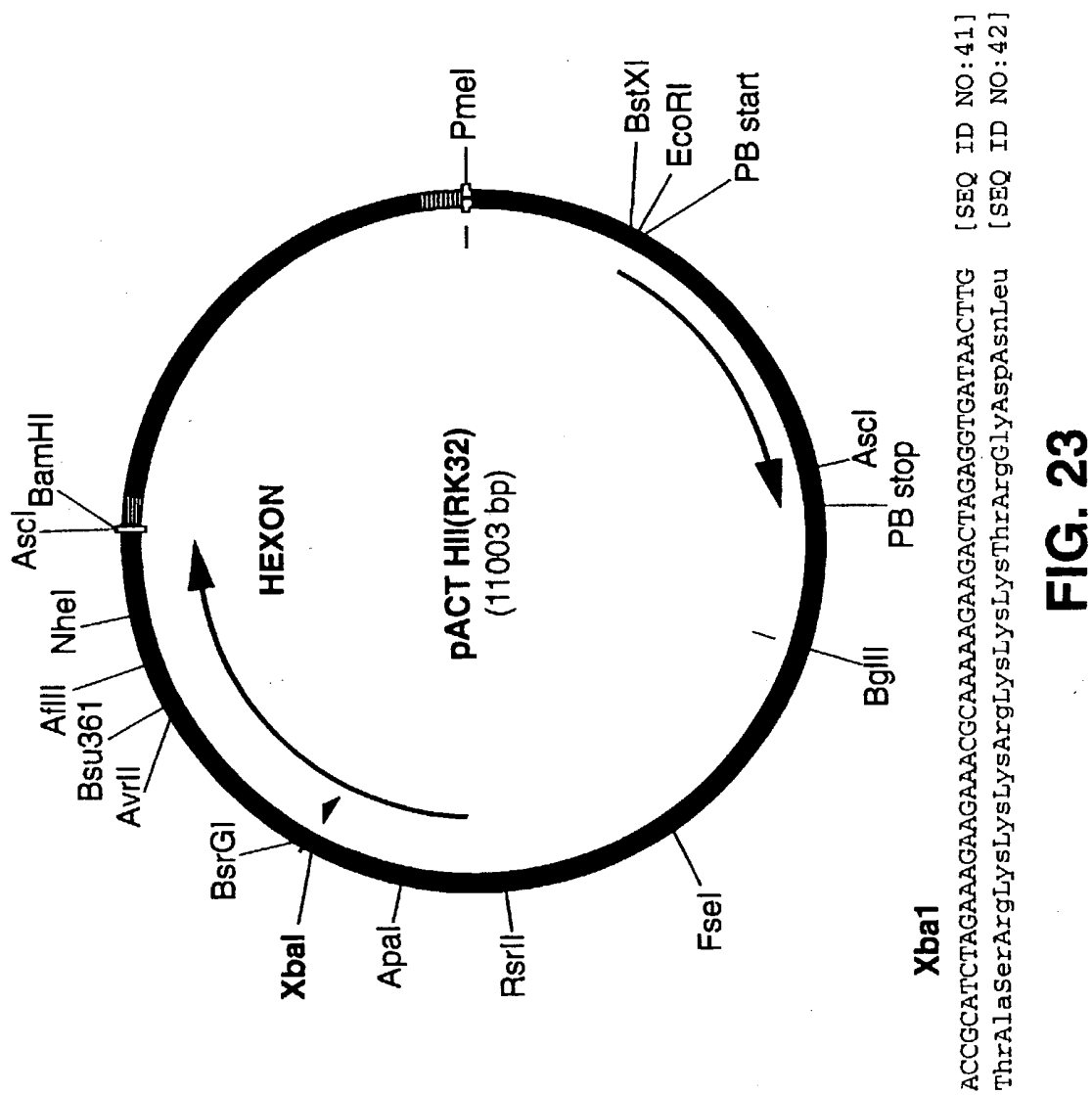
FIG. 21



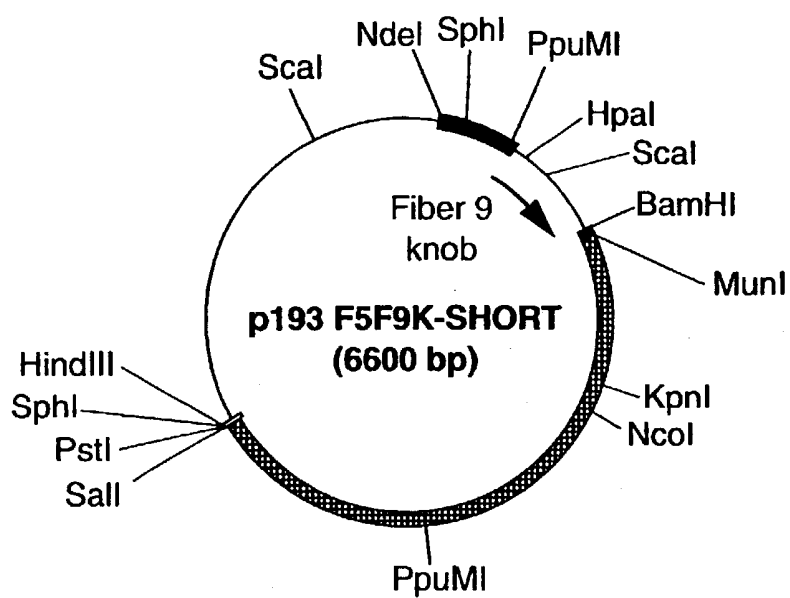
XbaI

ACCGCATCTAGAGGTGATAACTTG [SEQ ID No: 39]
 ThrAlaSerArgGlyAspAsnLeu [SEQ ID No: 40]

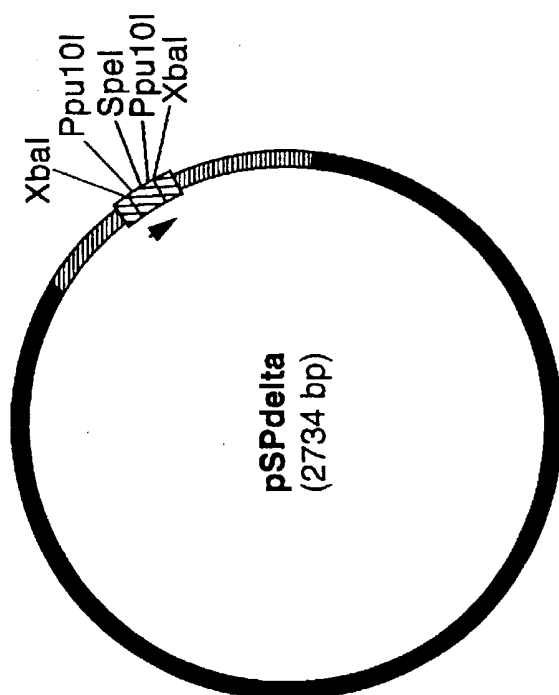
Figure 22



25/30

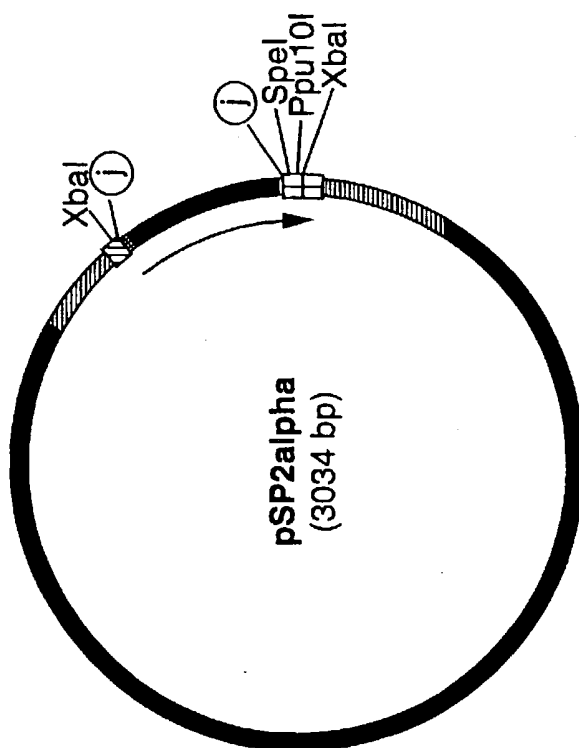
**FIG. 24**

26/30



XbaI Ppu10I SpeI Ppu10I XbaI
 TCTAGAGCAGCTATGTCATGAAGGACTAGTGGAGAGATGCATGCAGCCTCTAGA [SEQ ID NO:43]
 SerArgAlaAlaMetHisGluGlyThrSerGlyGluMetHisAlaAlaSerArg [SEQ ID NO:44]

FIG. 25



xbaI

TCTTAGAGCAGCTATGCAGGCGGCGCAGAAAGCTGAAGAGGCAGCCACACGGGCTGAGGAGAAAGCGCGCTGAGGCCGAAG
SerArgAlaAlaMetGlnAlaAlaAlaGluAlaGluAlaAlaThrArgAlaGluGluLysArgAlaGluAlaGluAla
CAGCGGCCGAAGCTGCCGCCCGCTGCGCAACCCGAGGTGCGAGAAGCCTCAGAAACAGAGCTAACGCTAAGGCCGA
laAlaAlaGluAlaAlaAlaProAlaAlaGlnProGluValGluLysProGlnLysGlnAlaAsnAlaLysAlaGln
AGCTGTGCAGGCGGCCGACAGAGTGAAGAGGCAGCCACACGGGCTGAGGAGAAAGCGCGCTGAGGCCGAAGCAGCGGCC
AlaAlaValGlnAlaAlaAlaGluAlaGluAlaAlaThrArgAlaGluGluLysArgAlaGluAlaGluAlaAlaAla

SpeI Ppu10I XbaI

ATGAAGGGACTAGTGGAGAGATGCATGCAGCCCTCTAGA [SEQ ID NO: 45]
isGluGlyThrSerGlyGluMetHisAlaAlaSerArg [SEQ ID NO: 46]

FIG. 26

28/30

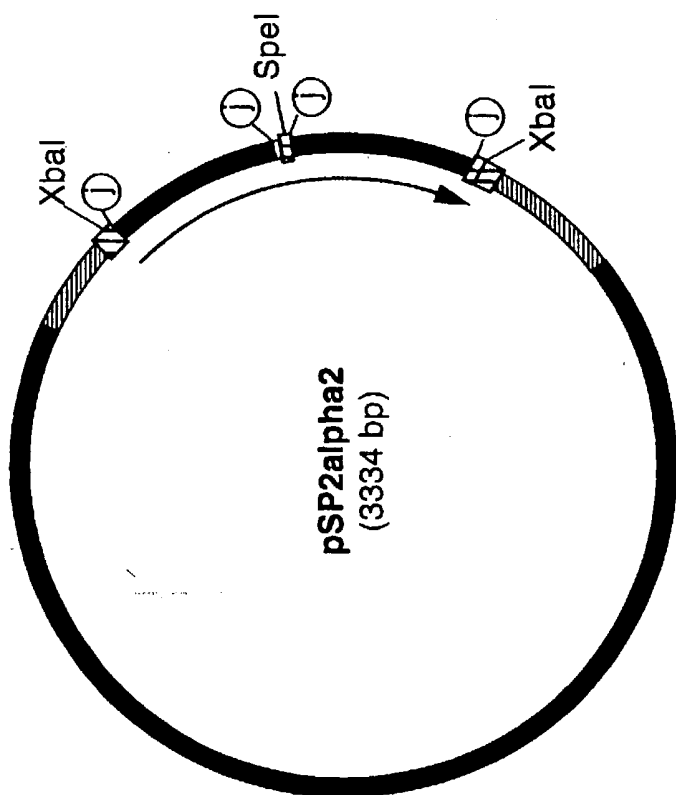


FIG. 27

29/30

XbaI ⑦

TCT AGA GCA GCT ATG CAG GCG GCC GCA GAA GCT GAA GAG GCA GCC
 Ser Arg Ala Ala Met Gln Ala Ala Ala Glu Ala Glu Glu Ala Ala
 ACA CGG GCT GAG GAG AAG CGC GCT GAG GCC GAA GCA GCG GCC GAA
 Thr Arg Ala Glu Glu Lys Arg Ala Glu Ala Glu Ala Ala Ala Glu
 GCT GCC GCC CCC GCT GCC CAA CCC GAG GTC GAG AAG CCT CAG AAG
 Ala Ala Ala Pro Ala Ala Gln Pro Glu Val Glu Lys Pro Gln Lys
 AAA CAG GCT AAC GCT AAG GCC GAA GCT GTG CAG GCG GCC GCA GAA
 Lys Gln Ala Asn Ala Lys Ala Glu Ala Val Gln Ala Ala Ala Glu
 GCT GAA GAG GCA GCC ACA CGG GCT GAG GAG AAG CGC GCT GAG GCC
 Ala Glu Glu Ala Ala Thr Arg Ala Glu Glu Lys Arg Ala Glu Ala
 GAA GCA GCG GCC GAA GCT GCC GCC CCC GCT GCG CAA CCC GAG GTC
 Glu Ala Ala Ala Glu Ala Ala Ala Pro Ala Ala Gln Pro Glu Val

⑦

GAG AAG CCT CAG AAG AAA CAG GCT AAC GCT AAG GCC GAA GCT GTG
 Glu Lys Pro Gln Lys Lys Gln Ala Asn Ala Lys Ala Glu Ala Val

SpeI ⑦

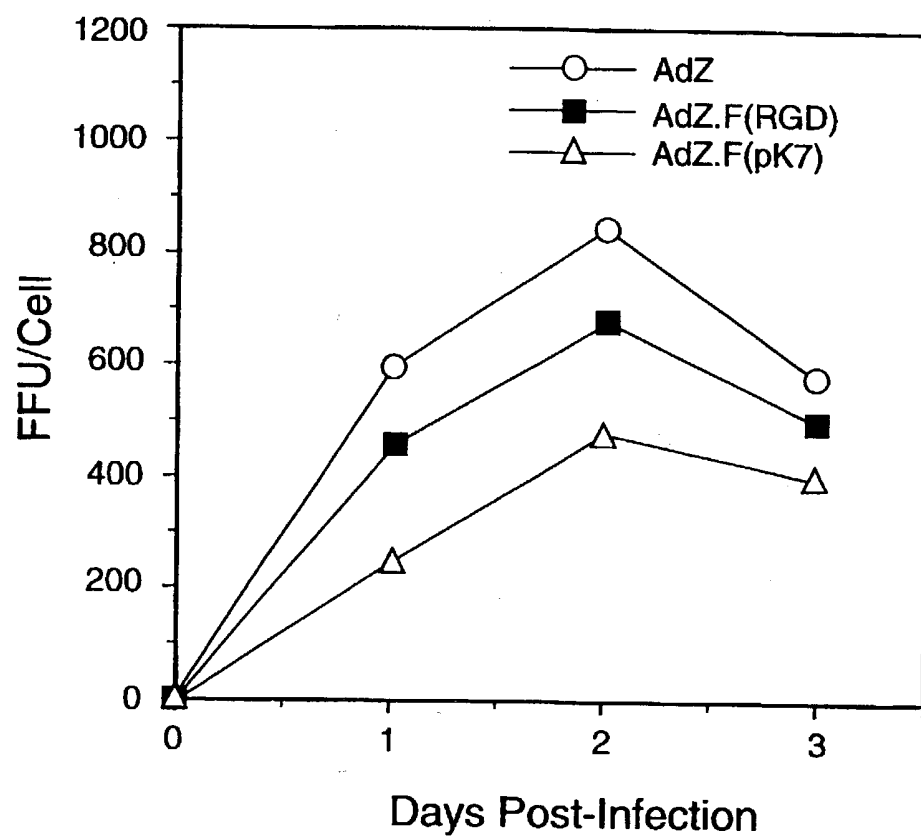
CAT GAA GGG ACT AGT GGA GAG ATG CAG GCG GCC GCA GAA GCT GAA
 His Glu Gly Thr Ser Gly Glu Met Gln Ala Ala Ala Glu Ala Glu
 GAG GCA GCC ACA CGG GCT GAG GAG AAG CGC GCT GAG GCC GAA GCA
 Glu Ala Ala Thr Arg Ala Glu Glu Lys Arg Ala Glu Ala Glu Ala
 GCG GCC GAA GCT GCC GCC CCC GCT GCG CAA CCC GAG GTC GAG AAG
 Ala Ala Glu Ala Ala Ala Pro Ala Ala Gln Pro Glu Val Glu Lys
 CCT CAG AAG AAA CAG GCT AAC GCT AAG GCC GAA GCT GTG CAG GCG
 Pro Gln Lys Lys Gln Ala Asn Ala Lys Ala Glu Ala Val Gln Ala
 GCC GCA GAA GCT GAA GAG GCA GCC ACA CGG GCT GAC GAG AAG CGC
 Ala Ala Glu Ala Glu Glu Ala Ala Thr Arg Ala Glu Glu Lys Arg
 GCT GAG GCC GAA GCA GCG GCC GAA GCT GCC GCC CCC GCT GCG CAA
 Ala Glu Ala Glu Ala Ala Ala Glu Ala Ala Ala Pro Ala Ala Gln
 CCC GAG GTC GAG AAG CCT CAG AAG AAA CAG GCT AAC GCT AAG GCC
 Pro Glu Val Glu Lys Pro Gln Lys Lys Gln Ala Asn Ala Lys Ala

⑦ **XbaI**

GAA GCT GTG CAT GCA GCC TCT AGA [SEQ ID NO:47]
 Glu Ala Val His Ala Ala Ser Arg [SEQ ID NO:48]

FIG. 27 Cont.

30/30

**FIG. 28**